Mathematische Kompetenzen von Kindergartenkindern

Überprüfung eines Testinstrumentes und Analyse von Unterschieden in der numerischen Leistungsentwicklung

Abhandlung
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der
Universität Zürich

vorgelegt von
Susanne Kuratli Geeler

Angenommen im Frühjahrsemester 2019
auf Antrag der Promotionskommission:

Prof. Dr. Elisabeth Moser Opitz (hauptverantwortliche Betreuungsperson)
Prof. Dr. Anke Lindmeier

Zürich, 2019

Danksagung

Ein besonderer Dank gebührt Prof. Dr. Elisabeth Moser Opitz für die wertvolle Unterstützung, die Anregungen und den fachlichen Rat bei der Umsetzung des vorliegenden Forschungsprojektes. Die Dissertation entstand im Rahmen des Projektes WILMA (Wir lernen Mathematik). Ein grosses Dankeschön gilt deshalb auch dem gesamten Projektteam für die gute Zusammenarbeit und Unterstützung, insbesondere auch Prof. Dr. Anke Lindmeier für die Begutachtung dieser Arbeit und Prof. Dr. Franziska Vogt, die stets ein offenes Ohr für meine Anliegen hatte. Danken möchte ich ebenfalls auch Dr. Urs Grob für die methodische Beratung.

Mein Dank richtet sich auch an alle Hilfskräfte, die für das Projekt im Einsatz waren, an die pädagogischen Fachkräfte und an die Kinder, die an dieser Studie teilgenommen haben.

Auch meinem Ehemann Christian und unsern Kindern Anna und Leon möchte ich herzlich danken, dass sie mich auf dem Weg von den Anfängen bis zur Fertigstellung dieser Arbeit begleitet, unterstützt und ermuntert haben.

Zusammenfassung

Eine Vielzahl von Forschungsergebnissen zeigt übereinstimmend, dass mathematische Kompetenzen von Kindergartenkindern wie *Zählen*, *Anzahl bestimmen* oder *Mengen vergleichen* zentrale Prädiktoren für spätere schulische Mathematikleistungen sind. Doch nur wenige Studien untersuchten bisher die Entwicklung mathematischer Kompetenzen im Kindergarten längsschnittlich über drei Testzeitpunkte. Die vorliegende Studie trägt zur Schliessung dieser Lücke bei, indem sie die mathematischen Leistungen von 894 Kindergartenkindern in der Schweiz und in Deutschland erfasst. Dabei stehen die Fragen nach Unterschieden in der mathematischen Entwicklung und der Einfluss diverser Kontextfaktoren im Vordergrund. Die Analyse der Daten erfolgte mittels Mehrebenen- und Wachstumskurvenmodellen. Die Ergebnisse bestätigen zum einen den aktuellen Forschungsstand, der von einer grossen Heterogenität in den frühen mathematischen Kompetenzen ausgeht. Dabei erklärt erwartungsgemäss das mathematische Vorwissen die Unterschiede in den Kompetenzen am besten, gefolgt vom kognitiven Leistungsvermögen. Zum anderen bilden die Ergebnisse die schnelle, aber auch unterschiedlich verlaufende, mathematische Leistungsentwicklung von Kindern zwischen vier und sechs Jahren ab. Der grösste Kompetenzzuwachs liess sich bei Kindern beobachten, die zum ersten Testzeitpunkt niedrige Leistungen gezeigt hatten. Im Gegensatz zu Ergebnissen aus der Grundschule nahm die Leistungsstreuung über die drei Testzeitpunkte im Kindergarten insgesamt nicht ab, sondern zu. Dieses Ergebnis weist darauf hin, dass sich im Kontext Kindergarten in Bezug auf die mathematische Kompetenzstreuung keine egalisierende Tendenz zeigt. Ein weiteres bemerkenswertes Ergebnis besteht darin, dass sich bei den Kindern in der Schweiz ein signifikant grösserer mathematischer Kompetenzzuwachs beobachten liess als bei den Kindern in Deutschland. Ob der höhere Leistungszuwachs auf Unterschiede im Kontext der frühen Bildung der beiden Länder zurückzuführen ist, wird diskutiert. Ebenso werden Ergebnisse betreffend unterschiedlicher Leistungsentwicklung von Mädchen und Jungen und Kindern mit Erstsprache Deutsch oder einer anderen Erstsprache analysiert und diskutiert. Es bedarf allerdings weitere Forschungsarbeiten, die mathematische Leistungen von Kindergartenkindern in unterschiedlichen Kontexten der frühen Bildung längsschnittlich messen und relevante Einflussfaktoren untersuchen.

Zugleich wird in dieser Studie überprüft, ob sich der statusdiagnostisch etablierte Test TEDI-MATH von Kaufmann et al. (2009) auch für die längsschnittliche Erfassung von mathematischen Kompetenzen im Kindergarten eignet. Anhand der Ergebnisse der Rasch-Analysen wird ersichtlich, dass ca. 40 % der Items einen schlechten Item-Fit aufweisen oder

von zeitbezogener Messvarianz bzw. von Messvarianz zwischen Subgruppen betroffen sind. Sie wurden deshalb aus den Analysen ausgeschlossen. Die verbliebenen 46 bzw. 54 Items zeigten sich indes Rasch-homogen und wiesen zu den drei Testzeitpunkten eine gute Reliabilität zwischen .92 und .94 auf. Die hohe Korrelation (r = .98) der Personenfähigkeitsparameter zwischen der Testversion mit allen Items und der reduzierten Version verdeutlicht, dass mit grosser Wahrscheinlichkeit auch mit wesentlich weniger Items dasselbe mathematische Konstrukt gemessen werden kann, was aus testökonomischer Sicht bedeutsam sein könnte und im Rahmen der vorliegenden Arbeit ebenfalls diskutiert wird.

Abstract

A large number of research results consistently show that the mathematical abilities of kindergarten children, such as *counting*, *counting how many* and *comparing quantities,* are important predictors of later educational performance in Mathematics. However, only a few studies have conducted longitudinal investigations of mathematical abilities in kindergarten children. The present study fills this gap by measuring the mathematical performance of 894 kindergarten children in Switzerland and Germany over three different testing times. The focus is on the questions of differences in mathematical development and the influence of various contextual factors. The analysis of the data was done using with multi-level and growth curve models. The results confirm, firstly, the current state of research, which assumes that there is significant heterogeneity in early mathematical abilities. As expected, previous mathematical knowledge was the strongest factor in explaining differences in abilities, followed by cognitive ability. Secondly, the results demonstrate the rapid development in mathematical performance in children between the ages of four and six years. The largest increase in ability was experienced by children who showed low performance at the first test time. In contrast to research results from primary schools, in this study, the distribution of competence increased. This result indicates that the kindergarten shows no equalizing tendency in terms of mathematical competence distribution. The children in Switzerland experienced a significantly greater increase in their mathematical ability than the children in Germany. Whether the higher performance gains are due to differences in the context of the early education of the two countries is discussed. Likewise, further results on different performance development of girls and boys and children with first language German and another first language will be analyzed and discussed.

At the same time, this study examines whether the TEDI-MATH test for status diagnosis developed by Kaufmann et al. (2009) is also suitable for the longitudinal study of the mathematical skills of kindergarten children. The results of the Rasch analyses showed that about 40% of the items displayed a poor item fit or were affected by variance in measurement over time or between sub-groups. These were excluded from the analyses. The remaining items proved to be Rasch homogeneous and displayed good reliability of between .92 und .94 at the three times of testing. The high correlation (r = .98) of the personal ability parameters between the test version with all items and the reduced version shows that it was possible to measure the same mathematical construct with considerably fewer items, which may be of significance for test efficiency and is likewise discussed.

Inhaltsverzeichnis

1 Einleitung

Im Zuge internationaler Vergleichsstudien wie TIMMS (Trends International Mathematics and Science Study) oder PISA (Programme for International Student Assessment) rückte auch die frühkindliche Bildung vermehrt in den Fokus von Politik, Wissenschaft und Gesellschaft. Mit dem von deutschen PISA-Studien erbrachten Nachweis, dass sich der Besuch eines Kindergartens positiv auf den Kompetenzerwerb in Mathematik auswirkt, erlangte die mathematische Förderung im Vorschulbereich vermehrte Aufmerksamkeit. Jugendliche, die in ihrer frühen Kindheit länger als ein Jahr eine Kindertagesstätte (KITA) besucht hatten, erzielten höhere mathematische Leistungen als Jugendliche, die keine oder weniger als ein Jahr eine KITA besuchten (Prenzel, Heidemeier, Ramm, Hohensee & Ehmke, 2004). In der Folge haben verschiedene nationale und internationale Forschungsergebnisse den Einfluss von KITA- bzw. Kindergarten-Besuchen auf diese frühen mathematischen Kompetenzen bestätigt und zudem aufzeigen können, dass mathematische Kompetenzen in der KITA bzw. im Kindergarten einen bedeutsamen Prädiktor für spätere Mathematikleistungen in der Schule darstellen (z. B. Gallit et al., 2018; Lyons, Price, Vaessen, Blomert & Ansari, 2014). Insbesondere das Zählen, die exakte Anzahlbestimmung (z. B. Toll, Kroesbergen & Van Luit, 2016) und das Vergleichen von Zahlen (z. B. Hornung, Schiltz, Brunner & Martin, 2014) sind von zentraler Bedeutung.

Die Entwicklung mathematischer Kompetenzen beginnt nicht erst mit dem Eintritt in einen Kindergarten. In sogenannten Habituationsexperimenten konnte nachgewiesen werden, dass bereits Säuglinge Unterschiede in numerischen Anordnungen von zwei bzw. drei Elementen erkennen (Wynn, 1992). Im Laufe der Zeit lernen die Kinder kontinuierlich hinzu. Sie erkennen, dass sich Mengen durch Hinzufügen oder Wegnehmen verändern lassen. Ungefähr ab dem zweiten Lebensjahr lernen sie einzelne Zahlwörter kennen (Fuson, 1988) und beginnen, Objekte zu zählen (z. B. Moser Opitz, 2008). Bei Kindergarteneintritt verfügen viele Kinder bereits über beachtliche mathematische Kompetenzen. Diese Kompetenzen sind allerdings sehr heterogen ausgeprägt und von verschiedenen Faktoren abhängig. Dazu tragen sowohl individuelle Merkmale wie Intelligenz (z. B. Hauser, Vogt, Stebler & Rechsteiner, 2014) oder Geschlecht (z. B. Winkelmann, Heuvel-Panhuizen & Robitzsch, 2008) als auch kontextuelle Faktoren wie sozio-ökonomischer Status (z. B. Schuchardt, Piekny, Grube & Mähler, 2014) oder Erstsprache (Anders, 2012) wesentlich bei.

Viele Studien überprüften bisher vor allem die Vorhersage schulischer Mathematikleistungen mittels mathematischer Vorläuferfertigkeiten im Kindergarten und den Einfluss von

Kontextfaktoren auf die Unterschiede in den mathematischen Leistungen in der Schule (z. B. Gallit et al., 2018; Jordan, Kaplan, Locuniak & Raminemi, 2007; Manfra et al., 2017). Nur wenige Studien untersuchten hingegen mathematische Kompetenzen im Kindergarten längsschnittlich über drei Testzeitpunkte und den Einfluss individueller und kontextueller Faktoren auf den Kompetenzzuwachs innerhalb des Kindergartens. Im Hinblick auf die Vorhersage von späteren mathematischen Kompetenzen scheint aber gerade dieser frühe mathematische Kompetenzzuwachs von zentraler Bedeutung zu sein. So zeigt eine Studie aus den USA mit über 1300 Kindern, dass der Leistungsfortschritt im Kindergarten der stärkste Prädiktor für die Mathematikleistungen von 15-Jährigen ist (Watts, Duncan, Siegler & Davis-Kean, 2014). In Bezug auf präventive Massnahmen ist es daher von Interesse, die Entwicklungsverläufe von Kindergartenkindern zu analysieren, um relevante Einflussfaktoren auf diese vorschulische mathematische Entwicklung eruieren zu können.

Die vorliegende Studie trägt zur Schliessung dieser Lücke bei und fokussiert die mathematische Entwicklung von 894 Kindergartenkindern über drei Testzeitpunkte. Dabei stehen Fragen nach Unterschieden in der mathematischen Entwicklung und der Einfluss diverser Kontextfaktoren wie allgemeine kognitive Fähigkeiten, Alter, Erstsprache oder Geschlecht auf die Leistungsentwicklung im Vordergrund. Zudem interessiert, wie sich Kinder mit niedrigem, mittlerem und hohem Vorwissen bezüglich ihrer mathematischen Kompetenzen entwickeln. Die Daten für diese Untersuchung stammen aus der vom Schweizerischen Nationalfonds und der Deutschen Forschungsgemeinschaft unterstützten WILMA-Studie der Universitäten Zürich und Landau-Koblenz, des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel und der Pädagogischen Hochschule St. Gallen (vgl. http://p3.snf.ch/project-156680). Das Design der WILMA-Studie war so angelegt, dass die Erhebungen in zwei Ländern durchgeführt werden konnten. Entsprechend gilt ein weiteres Forschungsinteresse vorliegender Untersuchung dem Vergleich mathematischer Kompetenzen und deren Entwicklung zwischen Kindern in Deutschland und Kindern in der Schweiz.

Um diese Fragestellungen beantworten und zuverlässige Aussagen machen zu können, ist ein Messinstrument notwendig, das die mathematischen Leistungen der Kindergartenkinder zu drei Testzeitpunkten reliabel und valide zu messen vermag. Längsschnittliche Untersuchungen und damit verbundene Veränderungsmessungen stellen generell hohe Anforderungen an die Messinstrumente, da sich der Messfehler entsprechend der Anzahl an Testungen erhöht und die Reliabilität sinkt (Ittel & Merkens, 2006). Zudem muss das Testinstrument der Entwicklung von Kindern dieser Altersstufe gerecht werden und über genügend Items im hohen sowie im niedrigen Leistungsbereich verfügen, um einen Decken- bzw. Bodeneffekt zu vermeiden (Döring & Bortz, 2016). Neben der mathematischen Kompetenzentwicklung von Kindergartenkindern unter verschiedenen Bedingungen gilt deshalb ein weiteres Forschungsinteresse dieser Studie der Überprüfung eines statusdiagnostisch etablierten Testinstrumentes, des TEDI-MATH von Kaufmann et al. (2009), zur Erfassung numerisch-rechnerischer Fertigkeiten. Von Interesse ist dahingehend, ob sich der Test auch zur längsschnittlichen, mathematischen Kompetenzerfassung von Kindergartenkindern eignet. In der klassischen Testtheorie (KTT) wird die Reliabilität eines Instrumentes üblicherweise mit Cronbachs Alpha angegeben. In Ergänzung dazu wird bei der Konstruktion neuer Leistungs- und Intelligenztests zunehmend das Rasch-Modell aus der Item-Response-Theorie (IRT) zur Prüfung der Instrumente beigezogen (Bond & Fox, 2015). In dieser Untersuchung wird der eingesetzte Mathematiktest deshalb auch mit dem Rasch-Modell auf Personen- und Itemhomogenität überprüft.

Den Ausgangspunkt vorliegender Arbeit bilden empirische Erkenntnisse und theoretische Hintergründe, die im ersten Teil der Arbeit in den Kapiteln 2 bis 6 dargestellt werden. Diese sind für den zweiten Teil der Arbeit, namentlich für die Hypothesenbildung der Fragestellungen und die Diskussion der Ergebnisse, wesentlich (Kapitel 7 bis 9).

Kapitel 2 widmet sich zunächst der frühen Bildung in Deutschland und der Schweiz, da ein zentrales Forschungsinteresse dieser Studie im Vergleich der mathematischen Kompetenzentwicklung zwischen Kindergartenkindern in Deutschland und der Schweiz liegt. Dabei werden Unterschiede und Gemeinsamkeiten bezüglich des frühen Bildungssystems, des Auftrags des Kindergartens und der Ausbildung von pädagogischen Fachkräften im Kindergarten aufgezeigt. Zusätzlich werden Forschungsergebnisse berichtet, die sich auf die mathematischen Kompetenzen und die Einstellung zum Fach Mathematik von pädagogischen Fachkräften beziehen.

Der Hauptfokus der gesamten Untersuchung liegt auf den mathematischen Kompetenzen von Kindergartenkindern. In Kapitel 3 wird deshalb die Bedeutung früher mathematischer Kompetenzen umrissen. Es werden Forschungsarbeiten aus der Schweiz und Deutschland sowie internationale Studien vorgestellt, die den Zusammenhang früher mathematischer Kompetenzen mit späteren Mathematikleistungen untersuchen. Abschliessend wird über Studien berichtet, die den Einfluss spezifischer mathematischer Kompetenzen auf spätere mathematische Leistungen überprüfen.

Da sich mathematische Kompetenzen nicht erst mit dem Eintritt in den Kindergarten entwickeln, widmet sich Kapitel 4 der Entwicklung dieser frühen mathematischen Kompetenzen ab dem Säuglingsalter. Dabei wird zwischen der approximativen und der exakten Mengenrepräsentation unterschieden und zwei aktuelle Modelle zum Zahlbegriffserwerb vorgestellt, in denen die Entwicklung früher Zahl-Mengen-Kompetenzen systematisiert sind.

In Kapitel 5 werden die Ergebnisse diverser Studien vorgestellt, die den Einfluss von Vorwissen, Intelligenz, Geschlecht, Alter und Merkmalen der sozialen Umwelt untersuchen. Vor diesem Hintergrund lassen sich schliesslich Hypothesen bezüglich des Einflusses von Kontextfaktoren auf die mathematischen Kompetenzen von Kindergartenkindern generieren.

Der erste Teil dieser Arbeit endet mit Kapitel 6, welches für die Beantwortung der Frage nach geeigneten Messinstrumenten zur längsschnittlichen Erfassung mathematischer Kompetenzen von Bedeutung ist. Es wird gezeigt, wie Ergebnisse von Testinstrumenten mit dem Rasch-Modell bezüglich Item-Fit, Subgruppeninvarianz und Messinvarianz über die Zeit geprüft und wie mit der Schätzung von Personenparameter valide Aussagen zur Kompetenzentwicklung gemacht werden können. Daneben werden, neben dem in dieser Untersuchung verwendeten TEDI-MATH (Kaufmann et al., 2009), drei weitere standardisierte und normierte Testinstrumente zur Erfassung mathematischer Kompetenzen im Kindergarten vorgestellt.

Mit Kapitel 7 beginnt der zweite Teil dieser Arbeit. Es handelt sich hierbei um die Darstellung der eigenen Untersuchung. Ausgehend von vier Forschungsfragen werden auf Basis des aufgearbeiteten Forschungsstandes und der theoretischen Auseinandersetzung Hypothesen gebildet. Danach werden das Untersuchungsdesign, die Stichprobe und die Erhebungsinstrumente vorgestellt. Darüber hinaus wird auf das methodische Vorgehen bezüglich Analysemethoden eingegangen.

Die Ergebnisse der Untersuchung werden in Kapitel 8 beschrieben. Zuerst werden hierzu die Ergebnisse der Überprüfung des mathematischen Testinstrumentes mit dem Rasch-Modell dargestellt. Danach werden die Ergebnisse zu den mathematischen Kompetenzen beim ersten Testzeitpunkt, zum Einfluss von Kontextfaktoren auf die Leistungsentwicklung zwischen den Testzeitpunkten und zum Entwicklungsverlauf von Kindern mit niedrigem, mittlerem und hohem Vorwissen beschrieben.

In Kapitel 9 werden zunächst die Ergebnisse dieser Studie zusammengefasst und die Hypothesen geprüft sowie diskutiert. Darauf folgt eine kritische Reflexion der Untersuchung, bevor abschliessend ein Fazit die wesentlichen Erkenntnisse festhält.

2 Vorschulische Bildung in Deutschland und der Schweiz

Internationale Vergleichsstudien zeigen, dass Schülerinnen und Schüler in verschiedenen Ländern über unterschiedliche mathematische Kompetenzen verfügen. So wiesen beispielsweise bei PISA 2015 Jugendliche aus den beiden asiatischen Ländern Japan und Korea durchschnittlich die höchsten mathematischen Leistungen auf, gefolgt von der Schweiz. Deutschland lag mit dem 11. Platz klar über dem OECD-Durchschnitt, aber deutlich hinter der Schweiz (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016). In der Literatur werden verschiedene Gründe für diese divergierenden Leistungen diskutiert. Unter anderem werden Unterschiede in der Lehrer- und Lehrerinnenausbildung (Blömeke, Kaiser, Döhrmann & Lehmann, 2010), in den fachlichen und fachdidaktischen Kompetenzen (Knievel, Lindmeier & Heinze, 2015) sowie in den Einstellungen von pädagogischen Fachkräften zu Mathematik (Li, Liu, DeBey, McFadden & Pan, 2018) als mögliche Ursachen diskutiert. Daneben werden auch kulturelle Hintergründe hinsichtlich der Unterschiede in der mathematischen Kompetenz aufgeführt. Mit Blick auf den Kindergarten konnten beispielsweise Huntsinger, Jose, Liaw und Ching (1997) in ihrer Forschung aufzeigen, dass chinesisch-amerikanische und taiwanesisch-amerikanische Kinder signifikant höhere Kompetenzen im Kindergarten besassen als europäisch-amerikanische Kinder. Mit dem Nachweis, dass sich die Dauer des Kindergartenbesuches positiv auf die mathematischen Leistungen der Kinder mit 15 Jahren auswirkt (Prenzel et al., 2004), wurden auch Unterschiede in der frühen Bildung als Erklärung für unterschiedliche Leistungen in Betracht gezogen (Gasteiger, Brunner & Chen, 2018). Diese werden im Folgenden für die beiden Länder Deutschland und Schweiz erörtert, da in der vorliegenden Studie mathematische Kompetenzen von Kindergartenkindern aus beiden Ländern erhoben werden. Zuerst wird ein Vergleich zwischen der Schweiz und Deutschland bezüglich der Institution Kindergarten, als Teil des frühen Bildungssystems, gezogen. Ein historischer Abriss über die Entstehung und Entwicklung des Kindergartens erklärt im Anschluss diese heutigen Gemeinsamkeiten und Unterschiede. Mit Bezug auf die Diskussion von möglichen Ursachen für unterschiedliche mathematische Leistungen folgt danach ein Vergleich der Ausbildung zur frühpädagogischen Fachkraft in Deutschland mit der Ausbildung zur Kindergärtnerin in der Schweiz. Ebenso werden die mathematischen Bildungspläne im Kindergarten und die Kompetenzen und Einstellungen der Fachkräfte beider Länder beschrieben.

2.1 Unterschiede und Gemeinsamkeiten zwischen Deutschland und der Schweiz im frühen Bildungssystem

In Deutschland wird unter Kindergarten im Allgemeinen eine Tageseinrichtung zur Erziehung und Betreuung von Kindern ab dem zweiten oder dritten Lebensjahr bis zum Eintritt in die Schule verstanden. Der Begriff KITA steht für Kindertagesstätte und wird synonym zum Begriff Kindergarten verwendet. Die Eltern zahlen Elternbeiträge, die je nach Bundesland oder Träger unterschiedlich hoch sind. In der Schweiz besuchen die Kinder ab dem vierten Lebensjahr in der Regel für zwei Jahre den Kindergarten, wobei dieser keine Tageseinrichtung darstellt und unentgeltlich ist. In der Schweiz ist der Besuch des Kindergartens unterdessen in nahezu allen Kantonen obligatorisch. In 17 Kantonen besuchen alle Kinder den Kindergarten während zwei Jahren. In 7 Kantonen ist der Kindergarten für mindestens ein Jahr für alle Kinder obligatorisch. Nur im Kanton Graubünden ist der Kindergartenbesuch (noch) freiwillig. Aber auch hier besuchen 98 % der Kinder den Kindergarten (EDK, 2017). Während der Kindergarten in der Schweiz einen im Lehrplan festgehaltenen Bildungsauftrag hat und einen Teil der obligatorischen Schule darstellt, gehört in Deutschland die sogenannte Elementarbildung zum Sozialbereich (Oberhuemer, Schreyer & Neuman, 2010). Mit dieser sozialpädagogischen Ausrichtung ist primär ein Erziehungs- und Betreuungsauftrag verbunden. Im Gegensatz zur obligatorischen Schule hat der Staat in der Kindertagesbetreuung keinen vom Erziehungsrecht der Eltern unabhängigen Auftrag, wie dies in der Schweiz der Fall ist. Entsprechend gibt es in Deutschland viele private und eher wenige öffentliche Träger, in der Schweiz dagegen vor allem öffentliche und nur sehr wenige private Träger (Gasteiger et al., 2018). Eine Verpflichtung zum Besuch eines Kindergartens besteht in keinem Bundesland. Trotzdem besuchen auch in Deutschland 93.6 % aller Kinder im Alter von drei bis fünf Jahren eine Kindertagesbetreuung (DESTATIS, 2017). Die wöchentliche Zeit, die die Kinder im Kindergarten verbringen, ist in Deutschland abhängig von den Buchungszeiten. In der Schweiz werden je nach Kanton im ersten Kindergarten zwischen 16 und 20 Lektionen und im zweiten Kindergartenjahr zwischen 22 und 25 Lektionen (EDK, 2017) im Kindergarten verbracht.

Der nachfolgende historische Abriss über die Entwicklung des Kindergartens von dessen Gründung Ende des 18. Jahrhunderts bis heute trägt zur Erklärung dieser Unterschiede bei.

2.1.1 Entstehung und Entwicklung der Kindergärten

Entstehung der ersten Kindergärten im 18. Jahrhundert

Im europäischen Raum lag die Erziehung von Kindern bis zum Alter von sechs/sieben Jahren historisch gesehen in der Verantwortung der Familie. Als Folge der Industrialisierung gegen Ende des 18. Jahrhunderts und der in Arbeiterfamilien vorherrschenden Armut mussten sowohl Väter als auch Mütter sowie ältere Kinder in Fabriken arbeiten. Dadurch konnte die Betreuung kleiner Kinder in den Familien nicht mehr sichergestellt werden. Es entstanden erstmals ausserfamiliäre Betreuungsinstitutionen, sogenannte Kleinkinderschulen (Konrad, 2004). Dabei waren zwei unterschiedliche Konzepte vorherrschend. Nach dem Schweizer Pädagogen Johann Heinrich Pestalozzi (1746–1827) sollte die Kleinkinderschule eine Nachbildung der familiären Wohnstube darstellen, wobei die Kinder nicht nur erzogen, sondern nach den Grundsätzen von Kopf (intellektueller Bildung), Herz (sittlicher Bildung) und Hand (physischer Bildung) auch unterrichtet wurden (Tenorth, 2010). Daneben gab es Konzepte, die die Auffassung vertraten, dass Kinder über Drill zu religiösen und gesellschaftsfähigen Personen zu erziehen seien (Berger, 2016). Diese früh gegründeten Kindergärten waren in erster Linie bürgerliche Einrichtungen, ohne einen sozialen oder religiösen Auftrag, wobei Kirchen, private Unternehmen und Gönner die Einrichtungen finanziell unterstützten. Damit konnten auch minderbemittelte Familien, die aufgrund ihrer Arbeitslage auf eine ausserfamiliäre Betreuung angewiesen waren, ihre Kinder in Kindergärten schicken (Lehnemann, 1995).

Entwicklung zwischen 1800 und 1900

In der ersten Hälfte des 19. Jahrhunderts entwickelte der deutsche Pädagoge Friedrich Fröbel (1782–1852) die Grundsätze Pestalozzis weiter. Die Ideen Pestalozzis lernte Fröbel 1805 an der Pestalozzi-Musterschule in Frankfurt am Main kennen. Er war fasziniert von den Schriften Pestalozzis. Von 1808 bis 1810 lebte er als Hauslehrer und Schüler Pestalozzis am Institut in Iferten in der Schweiz. Während dieser Zeit entdeckte er die besondere Bedeutung der frühen Kindheit in der menschlichen Entwicklung (Frey, Gehrlein & Wosnitza, 2006). Er entwickelte ein Konzept von Liedern, Bewegungsspielen und sogenannten Spielgaben, das in seinen gegründeten Kindergärten umgesetzt wurde. Als Spielgabe bezeichnete Fröbel ein Angebot von Materialien, insbesondere geometrischer Körper wie Ball, Kugel, Würfel, Walze und Kegel. Im Spiel mit diesen Materialien erschliesst sich dem Kind die Welt und ihre Gesetzmässigkeiten (Konrad, 2004). Fröbel gilt heute als eigentlicher Gründer des Kindergartens. Angesichts seiner methodischen Ansätze können Fröbels Kindergärten aus heutiger Sicht als erste Bildungsinstitutionen für junge Kinder betrachtet werden (Wullschleger, 2017). Dies führte dazu, dass die Kindergärten im Verlauf des 19. Jahrhunderts zunehmend weniger als soziales Angebot für Arbeiterfamilien, sondern vermehrt auch als Ergänzung zur familiären Erziehung betrachtet wurden. Die von Fröbel an Erziehung und Bildung orientierten Ansätze wurden anfänglich in Deutschland und insbesondere auch in der französischen Schweiz in den Kindergärten umgesetzt. Als Folge verankerte beispielsweise der Kanton Genf den Kindergarten bereits 1848 im Schulgesetz und erklärte ihn als Bestandteil des Schulwesens. Somit wurde der Kindergartenbesuch in Genf auch staatlich subventioniert (Nufer, 2012). Aus heutiger Sicht ist die Tatsache interessant, dass man sich damals in der Deutschschweiz skeptisch gegenüber dem Kindergarten als Betreuungs- und insbesondere als Bildungsinstitution zeigte. So wurde beispielsweise in den 1880er Jahren die Integration des Kindergartens in die Volksschule abgelehnt und somit eine „Verschulung" des Kindergartens verhindert (Walter-Laager, Joos & Fröhlich, 2013). Auch in Deutschland wurden im Verlauf des 19. Jahrhundert zunehmend Stimmen laut, die den Kindergarten als Bildungsinstitut nach dem Vorbild Fröbels ablehnten und vor allem ein Betreuungsangebot verlangten. Deutlich zeigte sich dies am Beispiel Preussens. Dort scheiterte 1851 der Versuch, den Kindergarten als elementare Form in das Bildungssystem zu integrieren und für Kinder, gleich welcher sozialen Herkunft, zu öffnen. Es kam sogar zu einem zehnjährigen Kindergartenverbot (Hemmerling, 2007). In der Folge blieb die sozialpädagogische Ausrichtung des Kindergartens in Deutschland bis heute bestehen.

Entwicklung zwischen 1900 und 1960

In der ersten Hälfte des 20. Jahrhunderts kamen mit der Reformpädagogik in beiden Ländern weitere Konzepte für den Kindergarten auf, die nicht selten in Konkurrenz zur Fröbelpädagogik standen. Zu erwähnen sind an dieser Stelle die Waldorfpädagogik, die auf der Grundlage der anthroposophischen Lehre von Rudolf Steiner (1861–1925) entwickelt wurde, und die Pädagogik der italienischen Ärztin und Pädagogin Maria Montessori (1870–1952). Während die Materialien Montessoris bis heute einen zielgerichteten und fachlichen Unterricht fokussieren, stellen Vertreter der Waldorfpädagogik vor allem die Anregung der kindlichen Fantasie und die natürliche Entwicklung in den Vordergrund (Wullschleger, 2017). Im Verlauf des 20. Jahrhunderts wurde der Kindergarten vor allem als Schonraum verstanden, der die Kinder einerseits vor der strengen Schule und andererseits auch vor dem wenig professionellen Erziehungsverständnis ihrer Eltern beschützen sollte (Witzig, 2013). Die Abgrenzung zur Schule wurde dabei in beiden Ländern stark hervorgehoben, was zur Folge hatte, dass der

Kindergarten im deutschsprachigen Raum zunehmend seine Bedeutung als Bildungsinstitution verlor und die fachliche Förderung in den Hintergrund geriet (Wittmann, 2010).

Entwicklung zwischen 1960 und 2000 in Deutschland

In den 1960er und 1970er Jahren war hinsichtlich der Aufgabe und Bedeutung des Kindergartens eine erneute Trendwende zu beobachten. Dies war einerseits den umfassenden und einflussreichen entwicklungspsychologischen Forschungsarbeiten von Jean Piaget zu den kognitiven Fähigkeiten von jüngeren Kindern und andererseits dem sogenannten Sputnik-Schock zu verdanken. Angesichts des Erfolgs russischer Wissenschaftler mit dem Satelliten „Sputnik“ rückte generell die mathematisch-naturwissenschaftliche Bildung in den Fokus und eine solche wurde zunehmend auch für den Kindergarten gefordert (Gasteiger, 2010). Trotzdem hatte in Deutschland die soziale Erziehung gegen Ende des 20. Jahrhunderts gegenüber der kognitiven Förderung Vorrang. Und obwohl der Kindegarten zu einem „unverzichtbaren Bestandteil des gesamten Bildungswesens“ (Deutscher Bildungsrat, 1970, S. 110) deklariert wurde, blieb er Teil des Sozialwesens. Dadurch wurde der sozialpädagogische Auftrag aufrechterhalten und der Kindergarten blieb in erster Linie ein ausserfamiliäres Betreuungsangebot. Damit verbunden ist bis heute eine Ganztagesbetreuung der Kinder und Kosten, die in Form von Elternbeiträgen zu leisten sind. Mit der Zuordnung des Kindergartens zum Bereich der Kinder- und Jugendhilfe (Oberhuemer et al., 2010) kann nicht vom Recht auf unentgeltliche Bildung Gebrauch gemacht werden.

Entwicklung zwischen 1960 und 2000 in der Schweiz

Anders verlief die Entwicklung des Kindergartens ab den 1960er Jahren in der Schweiz. Im Zuge der erwähnten bildungspolitischen Diskussionen wurde der Vorschulerziehung vermehrt Aufmerksamkeit geschenkt. Der Kindergarten wurde zunehmend als Bildungsinstitution wahrgenommen und der Aspekt der ausserfamiliären Betreuung geriet in den Hintergrund. Diese neue Sichtweise des Kindergartens führte dazu, dass auch vermehrt Kritik an den vermittelten Inhalten geübt wurde. Sie bezog sich insbesondere auf die zu wenig systematische kognitive Förderung, vor allem im Hinblick auf Kinder aus bildungsfernen Familien (Wannack, 2003). Daraufhin reagierte der Schweizerische Kindergärtnerinnenverein 1971 mit dem Erlass eines Rahmenplans für den Kindergarten, der von vielen Kantonen der Deutschschweiz als verbindlich erklärt wurde. Das Ziel des Rahmenplanes war es, „eine Übersicht über das Bildungsgeschehen im Kindergarten zu geben, sowie über Möglichkeiten und Methoden einer zeitgemässen Kindergartenführung zu informieren“ (Wannack, 2003, S. 26). Im Jahr 1997 wurde dieser im Auftrag der Erziehungsdirektion überarbeitet. Dem Kindergarten wurde neu ein Lehrauftrag erteilt, der sich zwar immer noch primär auf Erziehung bezog, aber doch auch fachliche Bildung miteinschloss.

Entwicklung ab 2000 in Deutschland

Eine erneute Trendwende löste in Deutschland die Veröffentlichung der ersten Ergebnisse der PISA-Studie 2000 aus. Die nachgewiesenen Bildungsdefizite 15-jähriger Jugendlicher im internationalen Vergleich lösten heftige Debatten aus. Und mit dem Nachweis, dass die Dauer des Kindergartenbesuchs einen Beitrag zur Vorhersage von Unterschieden in den mathematischen Kompetenzen liefert, rückte schliesslich auch die mathematische Bildung im Kindergarten in den Fokus der Diskussion (Prenzel et al., 2004). In der Folge wurden erste Bildungspläne erarbeitet (vgl. Abschnitt 2.1.2).

Entwicklung ab 2000 in der Schweiz

Analog zu Deutschland rückte auch in der Schweiz mit PISA 2000 die frühkindliche Bildung erneut in den Fokus von Politik und Wissenschaft. So empfahl die „PISA 2000 Steering Group" der EDK die Einführung einer Vorschulstufe, in welcher der „teilweise Übergang vom offenen Spiel zum systematischen Lernen anzustreben" (EDK, 2003, S. 16) sei. Diese Reform wurde 2006 schliesslich mit der Annahme des Bundesbeschlusses über eine Neuordnung der Verfassungsbestimmung zur Bildung durch das Volk eingeleitet. Der Bundesbeschluss beinhaltete insbesondere eine Koordination zwischen den Kantonen bezüglich einheitlicher Regelungen gewisser Eckpfeiler wie Schuleintrittsalter, Dauer und Ziele der Bildungsstufen, Anerkennung von Abschlüssen etc. (Bundesrat, 2006). Die Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) entwickelte in der Folge das HarmoS-Konkordat, das 2009 in Kraft trat und unter anderem die Einschulung kantonsübergreifend harmonisierte. Damit war der Kindergarten, als obligatorische zweijährige Bildungsinstitution, in praktisch allen Kantonen auch gesetzlich verankert (EDK, 2007).

2.1.2 Bildungs- bzw. Lehrpläne im Kindergarten

Bildungspläne in Deutschland

Die Bundesländer trafen im Jahr 2002, als Folge des PISA-Schockes, die Vereinbarung, Pläne für Bildung und Erziehung im Elementarbereich zu erstellen. Dazu wurde 2004 zuerst ein gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertagesstätten durch die Jugendministerkonferenz verfasst. Dieser sollte den einzelnen Bundesländern als Orientierung bei der Erstellung der Bildungspläne dienen. Betont wird darin, dass Lernen in Kindertageseinrichtungen im ganzheitlichen Sinne erfolgen soll, aber es geht auch hervor, dass der Kindergarten den Auftrag hat, die Bildungsmöglichkeiten des Kindes in den einzelnen Fachbereichen zu beachten und zu fördern (Jugendministerkonferenz, 2004). Für Mathematik gibt der Rahmenplan konkret vor, dass „die kindliche Neugier und der natürliche Entdeckungsdrang der Kinder dazu genutzt werden [sollten], den entwicklungsgemäßen Umgang mit Zahlen, Mengen und geometrischen Formen, mathematische Vorläuferkenntnisse und -Fähigkeiten zu erwerben" (Jugendministerkonferenz, 2004, S. 4). In der Folge entstanden in allen Bundesländern Bildungs- und Orientierungspläne für den Kindergarten, die sich sowohl bezüglich ihres Aufbaus als auch inhaltlich stark voneinander unterscheiden. Obwohl der Gesamtrahmenplan für das Fach Mathematik Hinweise der Förderung gibt, erwähnen nicht alle Bildungspläne der Bundesländer einen mathematischen Kompetenzbereich (Hasemann & Gasteiger, 2014). Da in dieser Studie Kinder aus den Bundesländern Schleswig-Holstein und Niedersachsen miteinbezogen wurden, werden diese beiden Pläne bezüglich ihrer mathematischen Orientierung genauer betrachtet. Im Bildungsplan Schleswig-Holsteins wird der Bildungsbereich Mathematik gemeinsam mit Naturwissenschaft und Technik aufgeführt. Zum Thema Mathematik heisst es: „Die Beschäftigung mit Zahlen und Grössen, mit Ordnen und Messen in Kindertageseinrichtungen kann Kindern die Welt mathematischer Zusammenhänge eröffnen" (Knauer & Hansen, 2012, S. 35). Danach werden Anregungen gegeben, wie Kinder mathematischen Themen begegnen können. Beispielsweise „Muster, Ornamente, Strukturen, Symmetrien begegnen den Kindern überall: in den Spuren, die der Regen im Sand hinterlässt, bei der Betrachtung eines Schneckengehäuses, im Muster der Tischdecke, beim Betrachten eines Kirchenfensters" (Knauer & Hansen, 2012, S. 35). Danach folgen in ähnlicher Weise Anregungen zum Sammeln, Vergleichen, Sortieren, zum Messen und Abwägen, Ideen um sich Raum und Zeit zu erschliessen oder Möglichkeiten, wie mathematische Alltagsaufgaben gelöst werden können.

Im Orientierungsplan für Kindertagesstätten in Niedersachsen wird im Kapitel „Mathematisches Grundverständnis" auf die Bedeutung mathematischer Tätigkeiten hingewiesen:

> „Mengen- und Größenvergleiche sowie (voroperationale) Tätigkeiten wie das Hinzufügen oder das Hinwegnehmen, das Aufteilen oder Verteilen sind fundamentale Handlungserfahrungen, mit denen viele Kinder nicht so vertraut sind. Daher sollten diese Aktivitäten in der Kindertagesstätte besonders in den Blick genommen und gezielt angebahnt werden.“ (Niedersächsisches Kultusministerium, 2011, S. 24)

Danach finden sich Beschreibungen und Anregungen von verschiedenen mathematischen Themen, die in den Alltag der Kindertagesstätte integriert werden können. Dabei handelt es sich um Ideen zum Sortieren, Klassifizieren und Quantifizieren von Grössen und Längen, von Raum-Lage-Beziehungen oder für das Raumerleben. Mit Augenmerk auf den Bereich Zahl und Mengen liest man, dass „die Mädchen und Jungen in unterschiedlichen Situationen im Alltag und im Spiel angeregt werden, Mengen zu erfassen und zu vergleichen. [...] Dabei wird mit zunehmendem Alter der Kinder auch das Zählen angebahnt und durch Spiele oder Abzählreime eingeübt.“ (Niedersächsisches Kultusministerium, 2011, S. 24–25).

Lehrplan in der Schweiz

Im Rahmen des HarmoS-Konkordats 2009 wurden als zentrales Instrument der Qualitätssicherung nationale Bildungstandards erarbeitet, worauf Lehrpläne, Lehrmittel und Evaluationsinstrumente abgestimmt werden (EDK, 2007). Diese Bildungsstandards wurden bei der Erarbeitung des neuen Deutschschweizerischen Lehrplans (Lehrplan 21) als Grundkompetenzen in verschiedenen Fachbereichen aufgenommen. Es werden nicht mehr Ziele für einzelne Schulstufen, sondern Kompetenzen für den jeweiligen Zyklus formuliert. Der Kindergarten gehört zum ersten Zyklus, der die beiden Kindergartenjahre und zwei Jahre Primarstufe umfasst. Zyklusübergreifend wird zuerst auf die Bedeutung des Fachs Mathematik, beispielsweise für die Umwelterschliessung, für die Entwicklung der Abstraktionsfähigkeit und Problemlösekompetenz oder für die Stärkung der Urteils- und Kritikfähigkeit, aufmerksam gemacht. Anschliessend werden allgemeine didaktische, strukturelle und inhaltliche Hinweise gegeben. Schliesslich folgen die drei Kompetenzbereiche des Fachs Mathematik: *Zahl und Variable*, *Form und Raum* und *Grössen, Funktionen, Daten, Zufall*. In allen Bereichen werden konkrete Kompetenzen beschrieben, die die Kinder im Verlauf der beiden Kindergartenjahre erwerben sollen. Da in der vorliegenden Studie die numerischen Kompetenzen im Vordergrund stehen, werden folgend nur die Kompetenzen aus dem Bereich Zahl und Variable vorgestellt (D-EDK, 2014):

Die Schülerinnen und Schüler...

- können Anzahlen vergleichen und die Begriffe grösser, kleiner, wird mehr, wird weniger, sind gleich viele, am meisten, am wenigsten verwenden,
- können bis zu 20 Elementen auszählen und im Zahlenraum bis 10 von jeder möglichen Zahl aus vor- und rückwärts zählen,
- können unterschiedliche Anzahlen einander angleichen (z. B. 8 und 4 Knöpfe → 6 und 6 Knöpfe),
- können Aussagen zu Anzahlen und Zahlpositionen an konkretem Material überprüfen (z. B. ein Turm mit 3 Klötzen ist höher als einer mit 2),
- können zeigen, wie sie zählen, und
- können Anzahlen verschieden darstellen (z. B. mit Punkten und Strichen) und verschieden anordnen (z. B. auf einer Linie und in der Fläche verteilt).

In der Schweiz werden im aktuellen Lehrplan 21, im Gegensatz zu früheren Lehr- bzw. Erziehungsplänen, keine Umsetzungshilfen oder Ideen der Förderung zur Zielerreichung mehr gegeben. Allen pädagogischen Fachkräften stehen aber verschiedene Lehrmittel, Materialien und Spiele zur Verfügung, die Möglichkeiten der mathematischen Förderung aufzeigen bzw. zur mathematischen Förderung eingesetzt werden können. Bis dato gibt es für den Fachbereich Mathematik im Kindergarten (noch) keine obligatorischen Lehrmittel, mit denen die Kindergärtnerinnen verbindlich zu arbeiten haben.

Vergleich der Bildungspläne zwischen Deutschland und der Schweiz

Vergleicht man den Lehrplan der Schweiz mit den beiden Bildungsplänen von Deutschland zeigt sich, dass in beiden Ländern jeweils auf die Bedeutung des Fachs Mathematik für die Erschliessung der Umwelt oder den Erwerb von Zusammenhängen etc. aufmerksam gemacht wird, wobei diese Beschreibungen im Lehrplan der Schweiz bedeutend mehr Platz einnehmen. Dies ist nachvollziehbar, da im Lehrplan in der Schweiz alle Lehrpersonen vom Kindergarten bis zum 9. Schuljahr und in den Bildungsplänen in Deutschland nur die pädagogischen Fachkräfte im Kindergarten angesprochen werden. Weiter fällt auf, dass in der Schweiz konkrete Kompetenzen beschrieben werden, die die Kinder im Kindergarten erwerben sollen. „Die Orientierungspunkte im ersten Zyklus geben an, an welchen Kompetenzstufen im Kindergarten verbindlich zu arbeiten ist“ (D-EDK, 2014, Gliederung der Volksschule). Dagegen werden in den Bildungsplänen von Niedersachsen und Schleswig-Holstein nicht verbindliche Kompetenzen oder Ziele formuliert, sondern mathematische Anregungen und

Ideen gegeben, die die pädagogischen Fachkräfte in den Kindergartenalltag einbauen können. Im Bildungsplan von Niedersachsen wird von einer „gezielten Anbahnung“ gesprochen, was auf etwas mehr Verbindlichkeit hinweist. Insgesamt fokussieren beide deutschen Bildungspläne eine alltagsintegrierte mathematische Förderung. Anregungen und Ideen werden immer im Zusammenhang mit Situationen im Spiel bzw. mit Handlungen, die sich in der Kindertagesstätte ergeben, geschildert. Im Gegensatz dazu findet die Förderung in der Schweiz nicht nur alltagsintegriert statt; neben den freien Spielphasen werden auch geführte Lektionen gehalten, in denen verschiedene Themen gezielt erarbeitet werden. Thematisch lassen sich in allen Bildungsplänen die Bereiche Zahlen, Mengen, Raum, Form und Grössen wiederfinden, wobei das Thema Zahlen und Mengen im Bildungsplan von Schleswig-Holstein eine eher untergeordnete Rolle spielt. Hier werden Anregungen und Ideen insbesondere zu Themen aus dem pränumerischen Bereich (Sortieren, Klassifizieren, Reihen bilden) und aus den mathematischen Bereichen Form und Raum bzw. Grössen beschrieben. Im Bereich Zahlen und Mengen sind nur wenig konkrete Hinweise zur Förderung enthalten. Es wird einzig darauf verwiesen, dass man Zahlen beim Tischdecken, Basteln und Planen brauche.

Insgesamt kann festgehalten werden, dass alle Bildungspläne mathematische Themen aufgreifen. Der Bildungsplan in der Schweiz gibt fachlich fundierte, verbindliche Kompetenzerwartungen in allen Bereichen der Mathematik vor, während die Bildungspläne in Deutschland zwar Anregungen zur mathematischen Förderung geben, darin aber wenig fachliche Systematisierung erkennen lassen (Fthenakis, 2009b).

2.1.3 Die Ausbildung zur pädagogischen Fachkraft im Kindergarten

Unterschiede zwischen den Ländern bestehen auch in der Ausbildung von pädagogischen Fachkräften, die für die Betreuung und Bildung in den Kindergärten verantwortlich sind. Im Folgenden verdeutlicht wiederum ein kurzer historischer Abriss über die Entstehung der Ausbildung zur pädagogischen Fachkraft die heutigen Unterschiede in den beiden Ländern. Danach werden aktuelle Ausbildungsinstitutionen und deren inhaltliche Themen skizziert und miteinander verglichen.

Die Ausbildung zur Erzieherin in Deutschland

Heute arbeiten in den deutschen Kindergärten nicht nur ausgebildete Erzieherinnen und Erzieher, sondern auch Diplom-Sozialpädagoginnen, Diplom-Sozialarbeiterinnen, Kinderpflegerinnen und Praktikantinnen mit zum Teil ganz unterschiedlichen Ausbildungen. Da aber die Berufsgruppe der Erzieherinnen mit 78 % den Hauptteil des KITA-Leitungspersonals stellt (DESTATIS, 2017), wird folgend diese Ausbildung geschichtlich aufgearbeitet und vorgestellt. Auf eine detaillierte Beschreibung der Ausbildung anderer Berufsgruppen wird verzichtet.

1839 begann Friedrich Fröbel in Deutschland mit mehrmonatigen Ausbildungskursen für Kindergärtnerinnen. Damit legte Fröbel den Stein für die Professionalisierung der Erzieherinnen. Er stellte pädagogische und sozialpflegerische Aufgaben in den Mittelpunkt und verlangte von den angehenden Erzieherinnen neben theoretischen Grund- und Fachkenntnissen auch praktisches Geschick und Einfühlungsvermögen (Erning, Neumann & Reyer, 1987). Neben diesen ausgebildeten Erzieherinnen arbeiteten aber weiterhin vor allem Ordensschwestern und Diakonissinnen in Kindergärten und Kleinkinderschulen. Später errichtete Ausbildungsstätten befanden sich dann auch überwiegend in kirchlicher Hand. Es gab keine einheitliche Struktur. Der jeweilige Träger setzte Dauer, Ziele, Inhalte und Methoden nach eigenem Interesse oder Einsatzfeld fest. 1885 lehnte der preussische Kultusminister eine staatliche Ausbildung mit der Begründung ab, dass eine gute Erzieherin vor allem über Gemüt und Takt verfügen müsse. Diese Eigenschaften könnten nicht durch eine staatlich finanzierte Ausbildung erworben werden (Jansen, 2010). Erst 1908 wurde die Kindergärtnerinnenausbildung schliesslich staatlich geregelt. So wurde ein einjähriger Fachkurs an einer Frauenschule angeboten. In der Weimarer Republik wurde 1928 eine zweijährige gemeinsame Ausbildung für Hortnerinnen und Erzieherinnen eingeführt, wobei nicht mehr primär Fröbels Gedankengut die Inhalte und Themen der Ausbildung prägten, sondern reformpädagogische Bewegungen und insbesondere der Ansatz von Maria Montessori (Erning et al., 1987). „Mit Beginn der nationalsozialistischen Herrschaft 1933 wurde die reformpädagogische Bewegung und somit auch die Montessori-Pädagogik in Deutschland niedergeschlagen und verboten. Bis 1945 dienten die Erziehung im Kindergarten und die daran gekoppelte Ausbildung des pädagogischen Personals ausschliesslich den Zwecken des nationalsozialistischen Regimes“ (Metzinger, 1993, S. 125). Nach dem Zweiten Weltkrieg knüpften die einzelnen Bundesländer im Wesentlichen an die Kindergartenpädagogik der Weimarer Republik an. Die Ausbildungsstruktur in der heutigen Form basiert im Wesentlichen auf Rahmenvereinbarungen der Kultusministerkonferenz von 1967. Die Ausbildung zur Kindergärtnerin und Hortnerin wurde mit der zur Jugend- und Heimerzieherin zu einer gemeinsamen Ausbildung zum *Staatlich anerkannten Erzieher* an Fachschulen für Sozialpädagogik zusammengefasst. Die Ausbildung zur Erzieherin bzw. zum Erzieher erfolgt heute über eine dreijährige Fachschule für Sozialpädagogik und ist auf oberem Sekundar- bzw. Postsekundarlevel angesiedelt (Oberhuemer et al., 2010). Im Wissen, dass sich die verschiedenen Fachschulen für Sozialpädagogik in Deutschland in ihren Lerninhalten und Modulen unterscheiden, werden hier stellvertretend die Lerninhalte für angehende Erzieherinnen des Regionalen Bildungszentrums in Kiel genannt (RBZ1, o.D.). Dies, weil in vorliegender Untersuchung vor allem Kindergärten aus dem Raum Kiel einbezogen waren. Der fachrichtungsbezogene Unterricht der Schule für Sozialpädagogik in Kiel umfasst dabei sechs Lernfelder: Berufliche Identität und professionelle Perspektive weiter entwickeln, Pädagogische Beziehungen gestalten und mit Gruppen pädagogisch arbeiten, Lebenswelten und Diversität wahrnehmen, Verstehen und Inklusion fördern, Sozialpädagogische Bildungsarbeit in den Bildungsbereichen professionell gestalten, Erziehungs- und Bildungspartnerschaften mit Eltern und Bezugspersonen gestalten sowie Übergänge unterstützen und schliesslich Institution und Team entwickeln sowie in Netzwerken kooperieren. Daneben gibt es einen fachrichtungsübergreifenden Unterricht mit den Themen Deutsch, Kommunikation mit Sprachförderung, Naturwissenschaft und Technik sowie Wirtschaft und Politik. Im Lernfeld 4 im Lehrplan von Schleswig-Holstein, *Sozialpädagogische Bildungsarbeit in den Bildungsbereichen professionell gestalten,* wird für den Ausbildungsgang zur Erzieherin auf die Bedeutung der Bildungsbereiche auf die Entwicklung von Kindern aufmerksam gemacht. Als Bildungsbereiche werden neben Spiel und Theater, Musik und Rhythmik, Ästhetik und Kunst, Sprache, Literacy und Medien, Religion, Gesellschaft und Ethik, Natur und Umwelt, Gesundheit und Ernährung auch Mathematik, Naturwissenschaften und Technik genannt. Es sollen «auf den Bildungsbereich bezogene fachspezifische und sozial-pädagogische Kompetenzen erworben werden. Beides ist didaktisch-methodisch miteinander zu verbinden» (Ministerium für Bildung und Wissenschaft des Landes Schleswig-Holstein, 2013, S. 53).

Die Ausbildung zur Kindergärtnerin in der Schweiz

Im Gegensatz zu Deutschland sind in der Schweiz nicht verschiedene Berufsgruppen in den Kindergärten vertreten. Heute arbeiten einheitlich spezifisch ausgebildete Kindergärtnerinnen in den Kindergärten.

Auch in der Schweiz legte Fröbel den Grundstein für die Professionalisierung der Kindergärtnerinnen. Da in der Schweiz noch keine Ausbildungskurse angeboten wurden, besuchten angehende Kindergärtnerinnen vorerst noch Ausbildungskurse in Deutschland. Es handelte sich um ehemalige Zöglinge des St. Galler Waisenhauses, die um 1870 am Fröbelinstitut in Nordhausen (D) ausgebildet wurden (Schlegel-Ganz, 2002). Das stetig wachsende Interesse junger Töchter in St. Gallen und in der ganzen Schweiz, den Kindergärtnerinnenberuf zu erlernen und die mit der Ausbildung in Deutschland verbundenen hohen Kosten, führten 1873 zur Gründung eines ersten einjährigen Ausbildungskurses in St. Gallen. Dieser bestand aus einem praktischen Teil (Falt-, Näh und Flechtschule etc.) und einem theoretischen (Fröbelsche Spielgaben, Pädagogik etc.). St. Gallen besass damit das erste Kindergärtnerinnenseminar der deutschsprachigen Schweiz (Schlegel-Ganz, 2002). Aufgrund des Erfolgs dieser ersten Ausbildungsstätte und der Nachfrage aus der ganzen Schweiz zur Aufnahme am Seminar entstanden im auslaufenden und beginnenden 19. Jahrhundert in verschiedenen Regionen der Schweiz kantonale und städtische Ausbildungsinstitute (Basel, Bern, Brugg, Luzern, Solothurn, St. Gallen und Zürich). Daneben gab es Kindergärtnerinnenseminare an katholischen Klosterschulen (Ingebohl, Menzingen, Baldegg, Cham) und ein evangelisches Kindergärtnerinnenseminar in Zürich (Schuh-Custer, 1969). Die damals einjährigen Kurse wurden ausgebaut, mit weiteren pädagogischen, psychologischen, handwerklichen und musischen Inhalten gefüllt sowie mit Praktika ergänzt und dauerten in der Mitte des 19. Jahrhunderts zwei (Schuh-Custer, 1969), gegen Ende des 19. Jahrhunderts durchschnittlich drei Jahre. 1995 beschloss die Schweizerische Konferenz der kantonalen Erziehungsdirektoren in ihren Empfehlungen zur Lehrerbildung, alle Lehrpersonen auf der Tertiärstufe auszubilden. Im Zuge dieser Umstrukturierung werden seit 2001 die zukünftigen Kindergärtnerinnen bzw. Kindergärtner nicht mehr in Kindergarten-Seminaren, sondern gemeinsam mit den Primarlehrpersonen an Pädagogischen Hochschulen ausgebildet. Mit dieser Reform wurde der Kindergarten nicht nur organisatorisch, sondern auch hinsichtlich des Personals und dessen Ausbildung an die Primarschule angebunden (Kucharz, 2014). Die Studierenden erwerben ein Bachelordiplom, das sie zum Unterrichten im Kindergarten und je nach Hochschule in der ersten und zweiten bzw. ersten bis dritten Klasse befähigt. Die Curricula der Hochschule legen den jeweiligen Ausbildungsanteil der einzelnen Disziplinen fest. Auf eine detaillierte Beschreibung der Curricula der verschiedenen Hochschulen wird verzichtet. Stellvertretend werden hier die Themen der Pädagogischen Hochschule St. Gallen aufgeführt, im Wissen, dass diese in den verschiedenen Hochschulen variieren können: Erziehungswissenschaften (38 ECTS), Sprachen (20 ECTS), Natur, Mensch und Gesellschaft und Mathematik (26 ECTS), Gestalten, Musik und Sport (38 ECTS), berufspraktische Studien (40 ECTS), Diverses (18 ECTS). In Bezug auf Mathematik existiert im ersten Semester des Studiums ein Modul „Einführung in die Mathematikdidaktik". Darin werden mathematische Basiskompetenzen im Kindergarten, Addition und Subtraktion in der ersten Klasse sowie Form und Raum im Anfangsunterricht behandelt (PHSG, o.D.). Parallel dazu werden in einer Vorlesung allgemein mathematikdidaktische Themen wie Mathekonferenzen, der Einsatz von Arbeitsmitteln oder Kompetenzbereiche und Handlungsaspekte des Lehrplans vermittelt. Im zweiten Semester besuchen die Studierenden das Modul „Fachdidaktische Grundlagen der Grundoperationen". Hier werden Grundvorstellungen, halbschriftliche und schriftliche Verfahren der Addition, Subtraktion, Multiplikation und Division erarbeitet. Im dritten Semester absolvieren die Studierenden das Modul „Mathematik in der Unterstufe". In diesem stehen für die Unterstufe (erste bis dritte Klasse) relevante Aspekte der Kombinatorik und der Wahrscheinlichkeit sowie die Themen Sachrechnen und Erarbeitung von Grössen im Zentrum. Im vierten Semester nehmen die Studierenden mit Diplomtyp A (Kindergarten und erste bis dritte Klasse) am Modul „Mathematik im Kindergarten" teil, in welchem vertiefend mathematische Themen des Kindergartens besprochen und verschiedene Konzepte der Vermittlung thematisiert werden.

Vergleich der Ausbildung zur pädagogischen Fachkraft im Kindergarten zwischen Deutschland und der Schweiz

In der Schweiz ist die Ausbildung zur Kindergärtnerin seit 2001 auf der tertiären Stufe, in Deutschland auf dem oberen Sekundar- bzw. Postsekundarlevel in Fachschulen für Sozialpädagogik, Fachakademien oder Berufsfachschulen angesiedelt (Oberhuemer et al., 2010). Obwohl in den letzten Jahren in Deutschland an Universitäten und Hochschulen zunehmend Bachelor- und Masterstudiengänge für die frühkindliche Bildung konzipiert wurden, ist der Anteil der pädagogischen Fachkräfte mit einem tertiären Bildungsabschluss mit 2.5 % relativ gering (DESTATIS, 2017). Die Ausbildungsinstitutionen *Pädagogische Hochschule* und *Fachschule für Sozialpädagogik* unterscheiden sich erheblich voneinander. Die Ausbildungsthemen und Schwerpunkte der Ausbildung sind kaum vergleichbar. In Bezug auf Mathematik werden in der Schweiz spezifisch fachdidaktische Themen in insgesamt vier Modulen erarbeitet, während in Deutschland der Lehrplan sehr allgemein für alle Bildungsfächer vorgibt, dass fachdidaktische und sozialpädagogische Kompetenzen zu erwerben sind.

2.2 Forschungsstand zur professionellen Kompetenz pädagogischer Fachkräfte im Kindergarten im Bereich Mathematik

In der Literatur werden von pädagogischen Fachkräften verschiedene professionelle Kompetenzen verlangt, die für das Lehren und Lernen von Bedeutung sind. Neben motivationalen und emotionalen Aspekten sowie selbstregulatorischen Fähigkeiten werden insbesondere das Professionswissen sowie professionelle Überzeugungen und Werthaltungen als relevant angesehen (Baumert & Kunter, 2006). Der Forschungsstand zu diesen beiden Kompetenzen wird nachfolgend aufgearbeitet.

Professionswissen

Zum Professionswissen werden in der Bildungsforschung zumeist das Fachwissen, das fachdidaktische und das allgemein pädagogische Wissen gezählt (Shulman, 1986). Unter Fachwissen werden das konzeptuelle Hintergrundwissen und das vertiefte Verständnis in einem Bildungsbereich verstanden. Das fachdidaktische Wissen bezieht sich auf Kenntnisse des Vermittelns von Inhalten des jeweiligen Faches, beispielsweise, wie Kinder mittels geeigneter Arbeitsmittel mathematische Kompetenzen erwerben. Allgemeines pädagogisches Wissen umfasst fachübergreifende Kenntnisse zur Gestaltung von Lernangeboten oder pädagogischen Interaktionen. Bisher existieren kaum empirische Befunde über die Bedeutung des Professionswissens von pädagogischen Fachkräften im Kindergarten (Anders, 2012). Trotz dieses Mangels wird jedoch eine teils lebhafte Diskussion über die Gewichtung von Fachwissen sowie pädagogischem und allgemein pädagogischem Wissen geführt (Anders, 2012). Insbesondere die Notwendigkeit von Fachwissen für pädagogische Fachkräfte im Kindergarten wird dabei unterschiedlich beurteilt. Es lassen sich diesbezüglich zwei Standpunkte ausmachen. Vertreter des Ko-Konstruktivismus argumentieren, „dass in einem pädagogischen Konzept, in dem sich die pädagogische Fachkraft gemeinsam mit dem Kind die Welt erschliesst, die Bedeutung des fachlichen Vorwissens für das pädagogische Handeln zu vernachlässigen sei“ (Anders, 2012, S. 29). Nach dieser Definition lernen Kinder durch die aktive Auseinandersetzung mit ihrer Umwelt und durch den Austausch mit anderen Kindern und Bezugspersonen. In Bezug auf den Kindergarten meint dieser Ansatz, dass Lernprozesse von Kindern und pädagogischen Fachkräften gemeinsam gestaltet werden. Kinder sind folglich aktive Konstrukteure ihres Wissens (Fthenakis, 2009a). Das Fachwissen der pädagogischen Fachkräfte spielt dabei eine untergeordnete Rolle. Vertreter des zweiten Standpunktes hingegen betonen die Bedeutung des Fachwissens für die pädagogische Qualität und weisen auf

Forschungsergebnisse hin. So wurden in einer Studie von Dunekacke, Jenßen und Blömeke (2015) 354 angehende Vorschullehrpersonen zu ihrem mathematischen Fachwissen mittels Papier-Bleistift-Test befragt. Daneben wurde mit Videovignetten überprüft, über welche Kompetenzen sie im Erkennen von mathematischem Potential in Alltagssituationen in der Kindertagesstätte verfügen und wie ihre Kompetenzen bezüglich der Planung von Situationen zur mathematischen Förderung gelagert sind. Die Ergebnisse der Studie legen nahe, dass das mathematische Fachwissen ein wichtiger Prädiktor für die Fähigkeit zur Wahrnehmung von mathematischen Lernsituationen und dies wiederum ein Prädiktor für die Planung von Fördersituationen ist. Dabei gab es auch einen signifikanten indirekten Effekt zwischen Fachwissen und der Planung von Fördersituationen. Der fehlende direkte Effekt wird durch das Autorenteam so interpretiert, dass die Qualität zur Planung von Lernsituationen davon abhängt, ob und wie angehende Vorschullehrpersonen mathematische Lernsituationen wahrnehmen. Da diese Fähigkeit wiederum vom Fachwissen abhängig ist, wird dessen Bedeutung bestätigt. Gleichzeitig zeigen die Ergebnisse, dass das Fachwissen zwar einen bedeutsamen Faktor darstellt, aber zur Planung von mathematischen Lernsituationen alleine nicht ausreicht.

Überzeugungen

Unter Überzeugungen werden subjektiv geprägte Vorstellungen zu verschiedenen Aspekten der beruflichen Tätigkeiten verstanden. Diese bedürfen weder einer Rechtfertigung, noch müssen sie widerspruchsfrei sein; der subjektive Glaube an die Richtigkeit genügt (Baumert & Kunter, 2006). Überzeugungen wird im Allgemeinen eine moderierende Wirkung auf die Wahrnehmung von Lehr-Lern-Situationen und das pädagogische Handeln zugesprochen. Sie haben einen Einfluss auf den Unterricht und auf das Lernen der Kinder (Pajares, 1992). Anders und Rossbach (2015) untersuchten in ihrer Studie den Einfluss von mathematischen Schulerfahrungen, emotionalen Einstellungen und pädagogischen Überzeugungen von 221 pädagogischen Fachkräften der Vorschule in Deutschland hinsichtlich der Sensibilität in der Erkennung von mathematischen Inhalten in spielbasierten Situationen. Dabei zeigte sich, dass insbesondere die gegenwärtige Freude und das Interesse an Mathematik die Sensibilität für das Erkennen von mathematischem Potential in Spielsituationen vorhersagten. Entgegen den Erwartungen der Forschergruppe beeinflussten eigene negative mathematische Schulerfahrungen die Anerkennung der Bedeutung von Mathematik als Bildungsbereich im Kindergarten nicht. Insgesamt liessen die pädagogischen Fachkräfte keine speziell positiven oder negativen Einstellungen gegenüber dem Fach Mathematik erkennen, sondern bewegten sich im neutralen Bereich.

Die Studie von Link, Vogt und Hauser (2017) verglich die Überzeugungen zur mathematischen Förderung von frühpädagogischen Fachkräften in Deutschland mit denen von Kindergärtnerinnen in der Schweiz. In beiden Ländern offenbarten die pädagogischen Fachkräfte eine deutliche Zustimmung zur konstruktionsorientierten Sicht auf das Lehren und Lernen von Mathematik. Pädagogische Fachkräfte mit einer konstruktionsorientierten Sicht stimmten dabei beispielsweise dem Finden unterschiedlicher Lösungswege durch die Auseinandersetzung mit mathematischen Inhalten zu oder befürworteten die selbstständige und kritische Diskussion mathematischer Probleme. In beiden Ländern zeigte sich eine dynamische Sichtweise von Mathematik. Dabei wird Mathematik als problembezogener Erkenntnis- und Verstehensprozess verstanden. Durch mathematische Tätigkeiten werden Sachverhalte verstanden oder Zusammenhänge erkannt (Post, Kastens & Lipowsky, 2013). Der Gegenpool wäre eine statische Sichtweise, die Mathematik als eine Sammlung von starren Rechenverfahren und Regeln auffasst, die zur Lösung von Aufgaben angewandt werden (Post et al., 2013).
Unterschiede zeigten sich indes bezüglich der Überzeugungen zur Gestaltung mathematischer Förderung im Kindergarten. Die pädagogischen Fachkräfte der Schweiz lehnten die Forderung, dass mathematische Förderung im Kindergartenalltag eher spontan stattfinden solle, ab. Unabhängig davon, ob sie in früheren Seminaren oder heutigen Hochschulen ausgebildet wurden, bevorzugten sie zielgerichtete und geplante mathematische Lernaktivitäten. Die pädagogischen Fachkräfte in Deutschland hingegen stimmten dem Konzept der spontanen mathematischen Förderung im Kindergarten eher zu und lehnten geplante Lernaktivitäten mehrheitlich ab. Daneben zeigten die pädagogischen Fachkräfte in der Schweiz eine deutlichere Ablehnung einer passiven Lernbegleitung im Spiel. Dies bringt zum Ausdruck, dass Kindergärtnerinnen in der Schweiz nicht davon ausgehen, dass Kinder mathematische Kompetenzen ohne Anregung durch die pädagogische Fachkraft von alleine erwerben würden. Die Autoren der Studie ziehen den Schluss, dass sich die Ergebnisse als ein „schulnäheres“ Verständnis der Aufgaben und der Rolle der frühpädagogischen Fachkraft in der Schweiz interpretieren lassen (Link et al., 2017).

Die Ergebnisse dieser Studie decken sich nur teilweise mit den Ergebnissen aus Forschungsarbeiten zu Überzeugungen bei deutschen Erzieherinnen im Elementarbereich. Benz (2012) und Levin, Wittmann und Bönig (2016) nutzten zur Erfassung von Überzeugungen zum Fach Mathematik unter anderem die Skalen Schemaorientierung und Prozessorientierung, die mit der in der Studie von Link et al. (2017) verwendeten statischen und dynamischen Sichtweise vergleichbar sind. Im Unterschied zu Link et al. (2017) gab es bei Benz (2012) eine deutliche Zustimmung der Fachkräfte zur Schemaorientierung (statische Sicht) und eine neutrale Haltung zur Prozessorientierung (dynamische Sicht). Insgesamt wurden 589 pädagogische Fachkräfte im Raum Karlsruhe befragt, wovon sich 308 noch in der Ausbildung befanden. Die Ergebnisse von Levin et al. (2016) sind vergleichbar mit der Studie von Link et al. (2017). Sie befragten insgesamt 769 Erzieherinnen in den Bundesländern Bremen und Baden-Württemberg unter anderem zur Überzeugung von Mathematik. Beobachten liess sich dabei eine eher ablehnende Haltung gegenüber der Schemaorientierung (statische Sicht) und eine deutliche Zustimmung zur Prozessorientierung (dynamische Sicht). Bei beiden deutschen Studien war, analog zur Untersuchung von Link et al. (2017), eine deutliche Zustimmung zu einer konstruktionsorientierten Sichtweise auf das Lehren und Lernen von Mathematik erkennbar.

2.3 Zusammenfassung

Der Kindergarten in der Schweiz und in Deutschland unterscheidet sich bezüglich verschiedener Aspekte erheblich. Historisch gesehen können in beiden Ländern die von Friedrich Fröbel gegründeten Kindergärten in der ersten Hälfte des 19. Jahrhunderts als Ursprung des heutigen Kindergartens betrachtet werden. Sie entstanden als Folge der fehlenden Betreuung von Arbeiterkindern in der Zeit der Industrialisierung und waren von den pädagogischen Grundsätzen Pestalozzis geprägt. Die weitere Entwicklung des Kindergartens verlief in beiden Ländern über eine lange Zeit in etwa gleich. Je nach Zeitgeist stand einmal das Betreuungsangebot und ein anderes Mal mehr das Erziehungs- und Bildungsangebot im Vordergrund. Erst mit der bildungspolitischen Diskussion rund um den Sputnik-Schock in den 1960er Jahren entwickelte sich der Kindergarten als Institution in den beiden Ländern unterschiedlich. In Deutschland blieb der Kindergarten, trotz Betonung der Bedeutung für die Bildung, dem Sozialwesen zugeordnet. Als Folge ist der Kindergarten bis heute in erster Linie ein ausserfamiliäres (kostenpflichtiges) Betreuungsangebot mit einem Erziehungsauftrag. In der Schweiz hingegen wurde als Folge der politischen Diskussionen in den 1960er Jahren der Kindergarten vermehrt als erste Bildungsinstitution wahrgenommen und näherte sich damit der Schule an. Heute gehört der Kindergarten in der Schweiz nahezu in allen Kantonen zur obligatorischen Schule und ist unentgeltlich. Es gibt verbindliche Lehrpläne mit klar formulierten Kompetenzen, auch im mathematischen Bereich, die die Kinder im Verlauf des

Kindergartens erwerben sollen. Obwohl im Zuge der Pisa-Forschung auch in Deutschland in praktisch allen Bundesländern Bildungspläne für den Kindergarten entstanden sind, spielt die mathematische Förderung eine untergeordnete Rolle und es existieren keine verbindlichen Ziele. Unterschiede bestehen darüber hinaus auch in der Ausbildung von pädagogischen Fachkräften für den Kindergarten. In der Schweiz erfolgt die Ausbildung zur Kindergärtnerin bzw. zum Kindergärtner seit 2001 auf Tertiärstufe an Pädagogischen Hochschulen gemeinsam mit den Primarlehrpersonen. In Deutschland werden die Erzieherinnen bzw. Erzieher in Fachschulen auf Sekundarstufe II ausgebildet. Zusammenfassend lässt sich festhalten, dass in den letzten Jahren in beiden Ländern Entwicklungen zur Annäherung von vorschulischen Institutionen an das schulische Bildungssystem stattgefunden haben. Dieser Prozess ist in der Schweiz aber deutlich weiter fortgeschritten.

Nachdem in diesem Kapitel die Gemeinsamkeiten und Unterschiede im frühen Bildungssystem in Deutschland und der Schweiz aufgezeigt wurden, liegt der Fokus in den nächsten Kapiteln 3 bis 5 auf den mathematischen Kompetenzen von Kindergartenkindern, ihrer Entwicklung und den Einflussfaktoren auf diese Entwicklung. Zunächst wird auf die Bedeutung früher mathematischer Kompetenzen für spätere Rechenleistungen eingegangen.

3 Bedeutung vorschulischer mathematischer Kompetenzen

Gemäss dem aktuellen Forschungsstand kann davon ausgegangen werden, dass mathematische Kompetenzen im Vorschulalter ein zentraler Prädiktor für die Entwicklung des arithmetischen Lernens in der Grundschule sind. Im Folgenden wird zuerst der Begriff *vorschulische mathematische Kompetenzen* definiert, bevor der Forschungsstand zum Zusammenhang von vorschulischen mathematischen Kompetenzen und späteren Mathematikleistungen aufgearbeitet wird. Der Abschluss des Kapitels präsentiert Ergebnisse aus Studien, die den spezifischen Einfluss unterschiedlicher vorschulischer mathematischer Kompetenzen auf spätere Mathematikleistungen fokussieren.

3.1 Begriffsklärung: vorschulische mathematische Kompetenzen

Der Kompetenzbegriff wurde in der Sozial- und Erziehungswissenschaft massgeblich von Emanuel Weinert geprägt. Weinert (2001) versteht unter Kompetenzen sowohl verfügbare als auch erlernbare Fähigkeiten und Fertigkeiten, um Aufgaben oder Lebenssituationen bewältigen zu können. Neben den kognitiven Fähigkeiten sind auch Motivation und Wille wesentliche Faktoren zur erfolgreichen Bewältigung. In der empirischen Bildungsforschung wird der Begriff hingegen enger gefasst. Es wird davon ausgegangen, dass Kompetenzen durch Erfahrungen und Lernen erworben werden können. Entsprechend sind sie auch durch äussere Interventionen beeinflussbar (Dorsch, Wirtz & Strohmer, 2017; Klieme & Leutner, 2006). Dabei können Kompetenzen selbst nicht unmittelbar gemessen werden, sondern sind nur durch die Performanz erschliessbar und werden zumeist leitungsbezogen erfasst. Entsprechend definieren Klieme und Leutner (2006) Kompetenz als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme & Leutner, 2006, S. 879). In vorliegender Arbeit werden Kompetenzen in diesem Sinn verstanden. Die Begriffe Kompetenz und Leistung bzw. Kompetenzentwicklung und Leistungsentwicklung werden synonym verwendet.

Der so verstandene Kompetenzbegriff wird durch den Zusatz *mathematisch* weiter spezifiziert. Die Kompetenzen beziehen sich folglich auf das Fachgebiet Mathematik. Dieses umfasst verschiedene Teilgebiete wie *Zahl und Variable*, *Form und Raum* oder *Grössen, Funktionen,* *Daten, Zufall* (D-EDK, 2014). In dieser Arbeit beziehen sich die mathematischen Kompetenzen auf den Bereich Zahl und Variable, wobei Zählen, Zahlen und Mengen im Vordergrund stehen. Unter *vorschulischen* mathematischen Kompetenzen werden schliesslich Kompetenzen verstanden, die Kinder im Bereich Zählen, Zahlen und Mengen vor dem Schuleintritt erwerben. Sie werden auch als mathematische Vorläuferfertigkeiten (z. B. Krajewski, 2008; Schneider, Küspert & Krajewski, 2016) mathematisches Vorwissen (z. B. Gasteiger, 2010) oder mathematische Basiskompetenzen (z. B. Krajewski & Ennemoser, 2013) bezeichnet. Hess (2012) hat insbesondere am Begriff der mathematischen Vorläuferfertigkeiten Kritik geübt und ihn durch mathematische Grundkompetenzen ersetzt. Dies unter anderem mit der Argumentation, dass jeder Kompetenz der Status der Vorläufigkeit inhärent sei, da sie sich entsprechend weiterentwickeln könne. In dieser Arbeit wird von vorschulischen oder frühen mathematischen Kompetenzen gesprochen.

3.2 Zusammenhang früher mathematischer Kompetenzen mit späteren Mathematikleistungen

Auf den Zusammenhang früher mathematischer Kompetenzen mit späteren Mathematikleistungen verweisen nicht nur Studien aus der Schweiz und Deutschland, sondern auch internationale Forschungsarbeiten. Im Folgenden werden zuerst Studien aus Deutschland und der Schweiz vorgestellt, bevor dann bedeutende Ergebnisse aus Luxemburg, Finnland, England und der USA berichtet werden. Es handelt sich bei allen Studien um Längsschnittstudien. Neuere Forschungsarbeiten werden jeweils zuerst erwähnt.

Studien aus Deutschland und der Schweiz

Gallit et al. (2018) untersuchten im Rahmen der SCHUES-Studie (Schulbezogene Umschriebene Entwicklungsstörungen – Prävention und Therapie unter Einbezug neuronaler Korrelate und des Entwicklungsverlaufs, 1. Phase) 1624 Kinder in Deutschland. Dabei erwies sich das Zahlen- und Mengenvorwissen (ZMW) im Kindergarten als bester Prädiktor für die Rechenleistungen in der ersten Klasse. Das ZMW wurde über Aufgaben zum Zählen (vorwärts und rückwärts), zur Grössen- und Mengenerfassung, zum Zahlenlesen sowie zur einfachen Addition und Subtraktion gemessen. Auch Dornheim (2008) konnte in einer Längsschnittuntersuchung mit 159 Kindern aus 14 Kindergärten aus sozial und regional verschiedenen deutschen Gegenden rund um Ludwigsburg nachweisen, dass das spezifische Zahlenvorwissen im Vorschulalter den Hauptprädiktor für die Rechenleistung in der ersten und zweiten Klasse der Grundschule darstellt. Die Testaufgaben beinhalteten Aufgaben zum Zählen, zur Anzahlerfassung sowie zum Zahlen lesen und schreiben. Daneben wurden auch das unspezifische Vorwissen (konzeptuelles Mengenverständnis und mathematikbezogenes Sprachverständnis) sowie allgemeine kognitive Fähigkeiten erhoben. In einer Längsschnittstudie aus der Schweiz (von Aster, Schweiter & Weinhold, 2007) wurde der Zusammenhang von mathematischen Grundkompetenzen im Kindergarten und späteren Rechenleistungen ebenfalls bestätigt. Diese untersuchten 337 Kinder aus dem Raum Zürich in urbanen und ländlichen Gebieten in der Schweiz. Die mathematischen Fertigkeiten wurden mit der neuropsychologischen Testbatterie für Zahlenverarbeitung und Rechnen (von Aster, 2001), der ZAREKI-K (T1, Kindergarten) und der ZAREKI-R (T2, zweite Klasse), erfasst. Eine deutsche, vierjährige Langzeitstudie mit 153 Kindern aus sechs Kindergärten rund um Würzburg ging der Frage nach, wie sich spezifische mathematische Basiskompetenzen im Kindergartenalter auf die Mathematikleistungen am Ende der ersten und vierten Klasse auswirken (Krajewski & Schneider, 2006). Dabei erklärten vorschulische Mengen-Zahlen-Kompetenzen (*Mengenvergleich* und *Seriation von Anzahlen*) am Ende der Grundschulzeit (vierte Klasse) 26 % der Varianz in den Mathematikleistungen. In einer weiteren Untersuchung mit 129 Kindern im Raum Freiburg ermöglichte das mathematische Vorwissen vor Schuleintritt eine genauere Vorhersage der zukünftigen schulischen Rechenleistungen als die nonverbale Intelligenz (Weisshaupt, Peucker & Wirtz, 2006). Die Erfassung des mathematischen Vorwissens erfolgte zu zwei Zeitpunkten (sechs und zwei Monate vor Schulbeginn) mit dem Diagnostikum zur Entwicklung des Zahlkonzepts (DEZ). Dieses Instrument erfasst mathematisches Vorwissen in zehn Bereichen. Die Rechenfertigkeiten der Kinder am Ende der ersten Klasse wurde mit dem DEMAT1+ (Krajewski, Küspert & Schneider, 2002) erhoben.

Internationale Studien

Manfra et al. (2017) erforschten den Zusammenhang von vorschulischen Fähigkeiten im Bereich Lesen und Mathematik mit den Fähigkeiten in der dritten Klasse von 1442 Kindern aus einkommensschwachen Familien in Florida (USA) aus verschiedenen ethnischen Gruppen. Die Ergebnisse legen nahe, dass die vorschulischen Kompetenzen generell ein starker Prädiktor für die Vorhersage der mathematischen Kompetenzen vier Jahre später waren. Insbesondere erwiesen sich in dieser Studie aber auch die frühen mathematischen Fähigkeiten als signifikanter Prädiktor für die Leistungsentwicklung; und dies unter Kontrolle verschiedener

Faktoren auf Kindergarten- und Schulebene. Die mathematischen Kompetenzen wurden mit Subskalen zum Zählen, Mengen und Objekte vergleichen, Objekte zählen sowie Zahlen in Reihe einordnen erhoben. Die Fähigkeiten in der dritten Klasse wurden mit dem high-stakes test (FCAT) gemessen. Dabei handelt es sich um einen landesweit üblichen Test zur sprachlichen und mathematischen Kompetenzmessung, der bei Schülerinnen und Schülern der dritten Klasse in Florida eingesetzt wird (Florida Department of Education, 2004). In einer weiteren, ebenfalls amerikanischen Studie mit einer umfangreichen Stichprobe von Watts et al. (2014) wurden die Daten von 1364 Kindern aus städtischen und ländlichen Gebieten der gesamten Vereinigten Staaten einbezogen. Das Autorenteam fand heraus, dass vorschulische mathematische Fähigkeiten die mathematischen Leistungen bis zum Alter von 15 Jahren voraussagen, wobei kognitive Fähigkeiten wie auch Familien- und Kindermerkmale kontrolliert wurden. Für die mathematische Leistungsmessung wurde ein Subtest des Woodcock Johnson-Revised (WJ-R) (Woodcock, McGrew & Mather, 2001) verwendet. Als ein noch stärkerer Prädiktor für Mathematikleistungen im Jugendalter erwies sich der Leistungsfortschritt der Kinder im Alter von 54 Monaten bis zur ersten Klasse. In einer weiteren amerikanischen Längsschnittstudie von Jordan et al. (2007) wurden die mathematischen Kompetenzen von 277 Vorschulkindern über insgesamt sechs Messzeitpunkte bis Mitte der ersten Klassenstufe erfasst und dann in Relation zur späteren mathematischen Leistungsentwicklung gesetzt. Gemessen wurden mathematische Kompetenzen wie Zählfertigkeiten, Zahlenkenntnis, Addieren und Subtrahieren. Als wesentlicher Befund liess sich festhalten, dass die mathematischen Vorkenntnisse von Kindergartenkindern 66 % der Varianz der Mathematikleistungen im ersten Schuljahr erklärten. Eine weitere interessante internationale Forschungsarbeit aus demselben Jahr stammt von Duncan et al. (2007). Hierbei handelt es sich um eine breit angelegte Studie, die Daten aus sechs Längsschnittstudien aus den USA, Kanada und England integriert. Es wurde auch hier ersichtlich, dass frühe mathematische Kompetenzen wie Zahlenkenntnis oder Zahlenreihenfolge die grössten Prädiktoren für spätere mathematische Leistungen waren.

Eine finnische Studie von Aunola, Leskinen, Lerkkanen und Nurmi (2004) konnte zum einen herausfinden, dass die Zählkompetenzen im Kindergarten die mathematischen Leistungen in der Schule am besten voraussagten. Zum anderen wurde deutlich, dass die mathematischen Fähigkeiten bei denjenigen Kindern, die bereits beim Eintritt in die Schule über höhere mathematische Kompetenzen verfügten, am schnellsten anstiegen. Untersucht wurden in dieser Studie 194 finnische Kinder im Kindergartenalter bis hin zur zweiten Klasse. Die mathematischen Kompetenzen wurden mit einem bestehenden Test zur Messung grundlegender mathematischer Konzepte (Ikaheimo, 1996) erhoben.

3.3 Einfluss von spezifischen Kompetenzen zur Anzahlbestimmung auf spätere Mathematikleistungen

Einige Forschungsarbeiten untersuchten zusätzlich, welches spezifische vorschulische mathematische Wissen besonders bedeutsam für die Vorhersage späterer Mathematikleistungen ist. Dabei wird zwischen frühen symbolischen versus nicht symbolischen Kompetenzen zur Anzahlbestimmung unterschieden (Kolkman, Kroesbergen & Leseman, 2013; Lyons et al., 2014; Toll et al., 2016). Das Autorenteam bezieht sich dabei auf zwei unterschiedliche Systeme zur Bestimmung von Anzahlen. Wenn Anzahlen durch Zählen exakt bestimmt und verglichen werden können, wird dies als symbolische Kompetenz bezeichnet. Das ungefähre Bestimmen bzw. Vergleichen von Anzahlen im Sinne von mehr oder weniger als nicht symbolisch (vgl. auch Kapitel 4.1). Hornung et al. (2014) sprechen in diesem Zusammenhang von „early number sense (ENC)“ und „nonverbal number senes“, Bonny und Lourenco (2013) von „symbolic number system“ und „approximate number system (ANS)“ und Obersteiner (2012) von „exakten Repräsentationen“ und „approximativen Repräsentationen“. Folgend werden die Ergebnisse dieser Längsschnittstudien aufgezeigt.

Toll et al. (2016) haben in einer Studie mit insgesamt 670 niederländischen Kindern festgestellt, dass die symbolischen Kompetenzen im Kindergarten (*Zählen*, *Bestimmen einer Anzahl Objekte mit und ohne Antippen*, *Rechnen mit Objektabbildungen*) neben der visuellen Arbeitsspeicherkapazität die mathematischen Leistungen in der ersten Klasse vorhersagen können. Sie waren sowohl für arithmetische Kompetenzen (*Addieren* und *Subtrahieren*) als auch für das Lösen mathematischer Problemstellungen im Sachkontext ein signifikanter Prädiktor. Hingegen war die nicht symbolische Kompetenz (*schneller Punktevergleich*) nur für die mathematischen Problemstellungen, nicht aber für den arithmetischen Faktenabruf in der ersten Klasse ein signifikanter Prädiktor.

In einer weiteren niederländischen Studie von Lyons et al. (2014) mit 1391 Kindern konnte gezeigt werden, dass das Wissen über Zahlbeziehungen und Zahlenreihenfolgen (symbolisches Wissen) in der ersten und zweiten Klasse insgesamt einen grösseren Beitrag zur

Varianzaufklärung der mathematischen Leistungsunterschiede in der sechsten Klasse leistete als das nicht symbolische Wissen. Aber auch eine nicht symbolische Kompetenz (*Mengenvergleich von Objekten)* in der ersten Klasse erwies sich als ein Prädiktor für die Vorhersage von Varianzunterschieden in der fünften Klasse. Hingegen leistete das schnelle Vergleichen von Punktemengen in allen Klassenstufen keinen Beitrag zur Varianzaufklärung. Diese Forschergruppe kam in einer späteren Untersuchung mit 539 Kindergartenkindern zu dem Schluss, dass die symbolischen Kompetenzen im Kindergarten die Erweiterung von nicht symbolischen Kompetenzen voraussagen (Lyons, Bugden, Zheng, De Jesus & Ansari, 2018).

Eine andere niederländische Studie von Kolkman et al. (2013) untersuchte 69 niederländisch sprechende Kinder über drei Testzeitpunkte im Alter von vier, fünf und sechs Jahren. Das Autorenteam fand erstens heraus, dass die nicht symbolischen Fähigkeiten (*Vergleich von Punktemengen* und *Punktemengen aufsteigend ordnen*) eine untergeordnete Rolle gegenüber den symbolischen Kompetenzen (*Zahlen kennen und benennen*) in der Entwicklung des Zahlverständnisses (*Zahlenreihenfolge* und *Zahlen vergleichen*) spielten. Zweitens erwiesen sich die symbolischen Aufgaben und die Aufgaben zum Zahlverständnis als wichtiger Prädiktor für die Vorhersage mathematischer Leistungen zum dritten Testzeitpunkt.

In einer luxemburgischen Studie (Hornung et al., 2014) wurde zwischen nichtverbalem Zahlverständnis (nonverbal number sense) und generell frühem Zahlverständnis (early number sense ENC) unterschieden. Die Aufgaben zum nichtverbalen Zahlverständnis sind vergleichbar mit den oben beschriebenen Aufgaben zur Erfassung der nicht symbolischen Kompetenz (*Mengenvergleich von Punkten* und *Objekten*). Ebenso lassen sich die Aufgaben zum ENC mit den Aufgaben zur Erfassung der symbolischen Kompetenz (*verbales Zählen*, *Anzahl bestimmen*, *Zahlen vergleichen*) vergleichen. Das nonverbale Zahlverständnis erwies sich dabei als Prädiktor für die Leistungen im frühen Zahlverständnis (ENC) im Kindergarten, während das frühe Zahlverständnis (ENC) die mathematischen Leistungen in der ersten Klasse vorhersagte. Insgesamt wurden in dieser Studie 165 Kinder über zwei Testzeitpunkte untersucht.

In einer amerikanischen Studie (Bonny & Lourenco, 2013) wurde der Einfluss nicht symbolischen Wissens (*Vergleichen von Punktemengen*) fokussiert. Insgesamt wurden 55 Kinder im Alter von drei, vier und fünf Jahren getestet. Es zeigte sich, dass der approximative Punktevergleich bei Kindern mit niedrigen mathematischen Kompetenzen einen Einfluss auf die frühen mathematischen Leistungen zum dritten Testzeitpunkt hatte. Kinder mit höheren nicht symbolischen Kompetenzen hatten dabei einen grösseren Kompetenzzuwachs. Für die Gesamtstichprobe liessen sich zwar auch Zusammenhänge zwischen nicht symbolischen Kompetenzen und späteren Rechenleistungen erkennen, allerdings waren diese nicht linear, sodass dahingehend keine Aussagen gemacht werden konnten.

Bei der Studie von Obersteiner (2012) handelt es sich um eine Interventionsstudie in ersten Klassen mit Längsschnittdesign. Insgesamt wurden 147 Kinder gleichmässig in vier Gruppen aufgeteilt. Die eine Gruppe erhielt eine Förderung zur approximativen Mengenbestimmung (*Mengenvergleich*, *Zahlvergleich*, *Überschlagsrechnen*), eine andere Gruppe zur exakten Mengenbestimmung (*Subitizing* und *Quasisimultanerfassung*) und eine weitere zu beiden Arten der Mengenbestimmung. Unter Subitizing wird das Erfassen von zwei bis maximal vier Gegenständen auf einen Blick verstanden, während Quasisimultanerfassung das Erkennen von grösseren Mengen auf einen Blick mit Hilfe einer strukturierten Darstellung (Würfelbilder, Zwanzigerfeld) meint (vgl. Abschnitt 4.1.1). Die Kontrollgruppe erhielt keine spezifische Förderung. Die Experimentalgruppen erzielten im Vergleich zur Kontrollgruppe signifikant höhere mathematische Leistungen. Es ergaben sich auch tendenziell höhere mathematische Leistungszuwächse, wenn entweder nur die exakte oder aber nur die approximative Kompetenz gefördert wurde. Die Kombination der beiden Ansätze führte tendenziell zu einer geringeren Kompetenzsteigerung.

In einer weiteren niederländischen Studie von Desoete, Ceulemans, De Weerdt und Pieters (2012) wurde zwischen den Bereichen „symbolic number world (NW)“, „symbolic arabic number (AN)“ und „non-symbolic comparison“ unterschieden. Es wurden folgende Aufgaben zur Erfassung dieser symbolischen Kompetenzen (NW und AN) eingesetzt: *Wörter als Zahlwörter identifizieren*, *Entscheiden, ob ein Zahlwort richtig ausgesprochen wurde*, *Die Länge von zwei gesprochenen Zahlwörter vergleichen* sowie *Zahlen von anderen Zeichen unterscheiden* und *Zahlen vergleichen*. Die Aufgabe zur Erfassung nicht symbolischer Kompetenzen war bei Desoete et al. (2012) übereinstimmend mit anderen Forschungsgruppen (Bonny & Lourenco, 2013; Kolkman et al., 2013; Lyons et al., 2014; Obersteiner, 2012; Toll et al., 2016) ein *Approximativer Punktevergleich*. Die Leistungen im Kindergarten bei der Lösung der symbolischen Aufgaben (NW und AN) waren ein Prädiktor für die arithmetischen Kompetenzen und den Faktenabruf in der zweiten Klasse, aber nicht für die arithmetischen Kompetenzen in der ersten Klasse. Die Leistungen der nicht symbolischen Aufgabe im Kindergarten sagten die Rechenleistungen in der ersten Klasse und den Faktenabruf in der zweiten Klasse voraus. Zudem kamen die Forschenden zu dem Schluss, dass die Kombination von Defiziten in den nicht symbolischen und den symbolischen Kompetenzen ein erhöhtes Risiko darstellt, dass sich eine mathematische Schwäche entwickelt.

3.4 Zusammenfassung

Insgesamt zeigen nationale und internationale Forschungsarbeiten auf eindrückliche Art und Weise die Bedeutung von mathematischen Kompetenzen im Kindergarten für spätere mathematische Leistungen. Die Kompetenzen im Kindergarten erwiesen sich durchwegs als bedeutendster Prädiktor zur Vorhersage mathematischer Leistungen in der Schule. Je höher die vorschulischen mathematischen Kompetenzen waren, desto höher waren tendenziell auch die späteren mathematischen Leistungen. In einer Studie wurde der mathematische Leistungsfortschritt im Kindergarten als der wichtigste Prädiktor in der Vorhersage der mathematischen Leistungen von 15-Jährigen identifiziert. Einige Forschungsarbeiten untersuchten, welches spezifische mathematische Vorwissen spätere Mathematikleistungen besonders gut vorhersagt. Dabei zeigten insbesondere Arbeiten aus den Niederlanden, dass frühe symbolische Kompetenzen die späteren Mathematikleistungen besser voraussagen können als nicht symbolische Kompetenzen. Unter symbolischen Kompetenzen werden in der Regel Aufgaben zum Zählen, zur Anzahlbestimmung, zum Lesen von Zahlen oder zum Vergleich von Zahlen verstanden. Im Gegensatz dazu spielten nicht symbolische Kompetenzen eine untergeordnete Rolle. Darunter werden Aufgaben zum ungefähren Bestimmen von Anzahlen (approximative Mengenbestimmung) oder zum ungefähren Vergleichen von Anzahlen (approximative Mengenunterscheidung) verstanden. Eine grosse Meta-Studie von Szkudlarek und Brannon (2017) bestätigte diese Befunde und zeigte zudem auf, dass der Einfluss von nicht symbolischen Kompetenzen tendenziell überschätzt wird (vgl. Lourenco & Bonny, 2017).

In diesem Kapitel wurde die Bedeutung mathematischer Kompetenzen im Kindergarten für spätere Mathematikleistungen beschrieben. Die Entwicklung dieser Kompetenzen beginnt aber nicht erst mit Beginn des Kindergartens, sondern bereits in frühster Kindheit. Im nächsten Kapitel wird diese Entwicklung aufgezeigt, bevor dann in Kapitel 5 auf Kontextfaktoren eingegangen wird, die die Entwicklung beeinflussen können.

4 Entwicklung mathematischer Kompetenzen

In Kapitel 3 wurde die Bedeutung früher mathematischer Kompetenzen für spätere Mathematikleistungen dargestellt. Dabei standen die mathematischen Fähigkeiten von Kindern im Alter zwischen vier und sechs Jahren und deren Vorhersage für spätere mathematische Leistungen im Zentrum. Heute ist bekannt, dass Kinder nicht erst mit Beginn des Kindergartens Kompetenzen im mathematischen Bereich erwerben (z. B. Schneider et al., 2016). So wie Kinder von Geburt an in verschiedenen Lebensbereichen lernen, entwickeln sich auch mathematische Konzepte und Kompetenzen ab dem Säuglingsalter.

In diesem Kapitel wird aufgezeigt, wie Kinder diese frühen mathematischen Kompetenzen erwerben. Dabei steht die Entwicklung im Bereich Mengen, Zählen und Zahlen im Zentrum. Es wird grundsätzlich zwischen zwei Zugängen zur Zahlbegriffsentwicklung unterschieden: approximative und exakte Mengenrepräsentation. Im Bereich der approximativen Mengenrepräsentation werden Erkenntnisse aus der Säuglingsforschung vorgestellt und die protoquantitativen Schemata nach Resnick (1989) erläutert. Im Bereich der exakten Mengenrepräsentationen stehen die Zählentwicklung nach Fuson (1988), die Zählprinzipien nach Gelman und Gallistel (1978) und verschiedene Strategien des zählenden Addierens und Subtrahierens im Zentrum. Diese verschiedenen Konzepte der Zahlbegriffsentwicklung werden in zwei aktuellen Modellen systematisch zusammengefasst. Diese werden im Anschluss vorgestellt und daraus Erkenntnisse für den Kindergarten gezogen. Dabei stehen die im Kindergarten zu erwerbenden mathematischen Kompetenzen und deren Stolpersteine im Vordergrund. Eine Zusammenfassung schliesst dieses Kapitel ab.

4.1 Zwei Zugänge zur Zahlbegriffsentwicklung

In der kognitiv-neurowissenschaftlichen Literatur wird eine Unterscheidung zwischen approximativer und exakter Mengenrepräsentation vorgenommen (z. B. Kucian, von Aster, Loenneker, Dietrich & Martin, 2008). Damit werden Prozesse, bei denen es um die ungefähre Repräsentation von Anzahlen geht, von solchen abgegrenzt, bei denen Mengen exakt repräsentiert werden müssen. Im Zusammenhang mit der Forschung zu frühen mathematischen Fähigkeiten wird auch von nicht symbolischen und symbolischen Kompetenzen gesprochen (vgl. Abschnitt 3.3). Unter nicht symbolischen Kompetenzen wird das ungefähre Bestimmen von Anzahlen verstanden, während mit symbolischen Kompetenzen das exakte Bestimmen einer Anzahl durch Zählen gemeint ist (z. B. Kolkman et al., 2013).

4.1.1 Approximative Mengenrepräsentation

Aus der Säuglingsforschung ist bekannt, dass bereits Säuglinge über approximative Mengenrepräsentationen verfügen (z. B. Antell & Keating, 1983; Starkey & Cooper, 1980; Xu, Spelke & Goddard, 2005). In sogenannten Habituationsexperimenten wurde die Aufmerksamkeitsspanne von Säuglingen über die Fixationsdauer des Blickes gemessen. Gewöhnen sie sich an einen gewissen Anblick, verringert sich die Blickdauer und bleibt in der Folge unverändert kurz. Die Präsentation bietet für die Kinder keinen neuen Reiz. Stellt der dargebotene Reiz aber eine Veränderung bzw. ein neues Ereignis dar, verlängert sich die Dauer der Fixation. Auf diese Weise konnten Starkey und Cooper (1980) bei durchschnittlich knapp sechs Monate alten Säuglingen nachweisen, dass diese kleine Anzahlen von zwei versus drei Punkten unterscheiden können. Bei einer grösseren Anzahl an Punkten (vier versus sechs und umgekehrt) gelang die Unterscheidung allerdings nicht. Dass Säuglinge zwischen kleinen Anzahlen unterscheiden können, die Differenzierung grösserer Anzahlen aber noch nicht gelingt, wurde in mehreren Untersuchungen repliziert (Antell & Keating, 1983; Strauss & Curtis, 1981; Wynn, 1996). Neuere Untersuchungen konnten nun aber belegen, dass Säuglinge auch Mengen mit mehr als drei Elementen unterscheiden konnten, allerdings nur, wenn die Differenz der Quantitäten gross genug war. So konnten Izard, Sann, Spelke und Streri (2009) Belege finden, wonach Neugeborene Anzahlen im Verhältnis 1:3 unterscheiden können, Anzahlen im Verhältnis 1:2 jedoch nicht. Sie schlossen daraus, dass gewisse Grundlagen numerischer Repräsentationen angeboren sind und sich mit zunehmendem Alter weiterentwickeln und ausdifferenzieren. In der Untersuchung von Xu, Spelke und Goddard (2005) konnten sechs Monate alte Kinder auch Anzahlen im Verhältnis 1:2 unterscheiden, nicht aber Anzahlen im Verhältnis 2:3. Dies war den Kindern durchschnittlich erst im Alter von neun bis zehn Monaten möglich. Anzahlen im Verhältnis 4:5 konnten allerdings noch nicht bewältigt werden (Xu & Arriaga, 2007). Auch wenn auf der Basis verschiedener Untersuchungen immer wieder Kritik an den Ergebnissen dieser Säuglingsforschung geübt wird (Mix, Levine & Huttenlocher, 1997; Moore, Benenson, Reznick, Peterson, & Kagan, 1987), herrscht Einigkeit darüber, dass bereits Säuglinge über ein kognitives Schema des Vergleichens von Mengen verfügen, wenn sich die Mengen in der Quantität deutlich voneinander abheben.

Nach Resnick (1989) gibt es neben dem Schema des Vergleichens noch zwei weitere Konzepte des approximativen Mengenverständnisses: das Schema des Vermehrens und Verminderns sowie das Teil-Ganzes-Schema. Diese drei Schemata werden von Resnick (1989) als protoquantitative Schemata bezeichnet.

Beim *Schema des Vergleichens* wird zwischen zwei Arten differenziert: Mengenunterscheidung mittels Simultanerfassung und Mengenunterscheidung mittels Schätzen. Bei Mengen bis drei oder vier Elementen kann die Differenzierung mittels Simultanerfassung, auch Subitizing genannt, erfolgen (Clements, 1999). Subitizing meint, dass Kinder zwei bis maximal vier Gegenstände auf einen Blick erkennen können, ohne dabei mathematisches Wissen zu nutzen. Diese Fähigkeit wird teilweise zur Erklärung der oben beschriebenen Säuglingsexperimente herangezogen (Starkey & Cooper, 1980). Grössere Anzahlen können auf einen Blick erfasst werden, wenn die Anzahl in kleinere Mengen zerlegt und zu neuen Einheiten zusammengefasst wird. Diesen Prozess der Wahrnehmung und Bestimmung der Anzahl wird als quasi-simultane Erfassung (Radatz, Schipper, Ebeling & Dröge, 1996) oder auch Conceptual Subitizing (Clements, 1999) bezeichnet. Kinder können aber auch grössere Mengen auf einen Blick erkennen, wenn diese in einer bekannten Struktur wie beispielsweise einem Würfel oder einem 20er-Feld dargestellt sind. Hier wird auch von strukturierter Anzahlwahrnehmung gesprochen (Benz, Peter-Koop & Grüßing, 2015). Bei grösseren unstrukturierten Mengen kann auf das System des Schätzens zurückgegriffen werden. Dieser Vorgang der Quantifizierung ohne genaue Anzahlbestimmung wird auch als approximate number system (Halberda & Feigenson, 2008) bezeichnet. Wie einleitend beschrieben, hängt die Unterscheidung von Mengen durch Schätzen sowohl von der absoluten Anzahl als auch von der Differenz der Anzahl zwischen den vergleichenden Mengen ab. Je älter die Kinder sind, desto eher gelingt das Unterscheiden von grösseren Mengen mit kleineren Differenzen. Kindergartenkinder im Alter von fünf bis sechs Jahren können in der Regel zwischen Mengen bis zu einem Verhältnis von 5:6 unterscheiden (Halberda & Feigenson, 2008). Mit zunehmender Grösse der Mengen wird die Unterscheidung immer schwieriger. Ist die Differenz zwischen zwei Mengen im Verhältnis zu ihrer Gesamtmenge zu gering, können sie nicht mehr unterschieden werden (Weisshaupt & Peucker, 2009).

Das *Schema des Vermehrens und Verminderns* bezeichnet die Einsicht, dass eine Menge durch Hinzufügen vergrössert und durch Wegnehmen verkleinert werden kann. Nach Resnick (1989) verfügen Kinder im Alter von drei bis vier Jahren über dieses Schema, wobei sie sich vermutlich auf die Kompetenzen auf verbaler Ebene beziehen und das zugrundeliegende Schema bereits deutlich früher vorhanden ist (Gerlach, 2007). Diese Vermutung wurde von Wynn (1992) in ihren Habituierungsstudien bestätigt. Sie stellte fest, dass bereits Säuglinge mit etwa fünf Monaten ein Verständnis von Additions- und Subtraktionshandlungen haben und erkennen, dass sich Mengen durch Hinzufügen bzw. Wegnehmen verändern lassen. Obwohl die Ergebnisse in mehreren Studien repliziert wurden (Berger, Tzur & Posner, 2006; Koechlin, Dehaene & Mehler, 1997; Simon, Hespos & Rochat, 1995), sind sie nicht unbestritten. So replizierten Feigenson, Carey und Spelke (2002) zwar Wynns Ergebnisse ebenfalls, stellten aber in einer Erweiterung des Versuchsaufbaus fest, dass die Kinder ihre Aufmerksamkeit eher auf Ausdehnungsmerkmale und weniger auf die Veränderung der Anzahlen richteten.

Das dritte protoquantitative Schema von Resnick (1989) wird als *Teil-Ganzes-Schema* bezeichnet. Kinder ab ca. vier Jahren verstehen, dass eine Menge in unterschiedliche Teilmengen zerlegt werden kann und dass verschiedene Teilmengen zu einer Gesamtmenge zusammengefügt werden können. Das Teil-Ganzes-Schema ist grundlegend für ein Verständnis von Zahlbeziehungen, Zahlzerlegungen und Rechenoperationen (Resnick, 1989).

Alle drei protoquantitativen Schemata sind zunächst nicht mit Sprache verbunden und werden nicht numerisch verstanden. Mit Beginn der Sprache und der Verknüpfung mit Zählfertigkeiten differenziert sich dieses Mengenverständnis weiter aus und ermöglicht den Erwerb eines tragfähigen Zahlbegriffs.

4.1.2 Exakte Mengenrepräsentation

Unter exakter Mengenrepräsentation wird das exakte Bestimmen einer Anzahl durch Zählen verstanden. Mit Beginn der sprachlichen Entwicklung lernen Kinder erste Zahlwörter kennen und ahmen diese nach. Anfänglich sind dies einzelne Zahlwörter oder kurze Zahlwortfolgen, die in ihrem Umfeld häufig genannt werden und mit einer bestimmten Bedeutung verbunden sind. Mit der Zeit erwerben sie nach und nach die exakte Zahlwortreihe. Nach Fuson (1988) läuft die Entwicklung der Zahlwortreihe in fünf Niveaus ab, die nachfolgend beschrieben werden. Die deutschen Begriffe beziehen sich auf Moser Opitz (2008).

1. *String level* (Ganzheitsauffassung der Zahlwortreihe)
Anfangs verstehen Kinder die Zahlwortreihe als undifferenziertes Wortganzes. Einzelne Zahlwörter werden noch nicht als solche erkannt. Die Kinder können eine kurze Sequenz von Zahlwörtern ähnlich einem Gedicht oder einem Lied aufsagen: „einszweidreivierfünfsechs". Die Zahlwortreihe kann in diesem Stadium jedoch noch nicht zum Zählen von Objekten verwendet werden. Eine Eins-zu-Eins-Zuordnung ist noch nicht möglich.

2. *Unbreakable list level* (unflexible Zahlwortreihe)
Die Zahlwörter werden in dieser Phase als eigenständige Einheiten erkannt und können von Nicht-Zahlwörtern unterschieden werden. Die Zahlwortreihe wird aber als „unzerbrechliche" Einheit wahrgenommen, die immer mit 1 beginnt. Entsprechend kann sie nur aufgesagt werden, wenn mit 1 begonnen wird. Es gelingt nicht, von einer beliebigen Zahl aus weiter zu zählen. Nachfolger und Vorgänger können nur bestimmt werden, indem das Kind sie innerhalb der Zahlwortreihe sucht. Anzahlbestimmungen durch Zählen sind auf diesem Niveau aber bereits möglich.

3. *Breakable chain level* (teilweise flexible Zahlwortreihe)
Kinder können nun bei jedem beliebigen Zahlwort mit dem Aufsagen der Zahlwortreihe beginnen. Nachfolger bzw. Vorgänger von Zahlen können bestimmt werden, ohne dass die gesamte Zahlwortreihe aufgesagt werden muss. Mit einer gewissen Verzögerung gelingt allmählich auch das Rückwärtszählen.

4. *Numerable chain level* (Flexible Zahlwortreihe)
Jedes Zahlwort wird als Einheit im numerischen Sinn verstanden. Die einzelnen Zahlwörter sind selbst zählbar. Dies ermöglicht, dass Kinder von einer bestimmten Zahl um eine vorgegebene Anzahl von Schritten weiterzählen können. Somit sind erste zählende additive Operationen möglich.

5. *Bidirectional chain level* (vollständig reversible Zahlwortreihe)
Es kann nun von jeder beliebigen Zahl aus vor- und rückwärts gezählt werden. Die Kinder auf diesem Niveau sind auch fähig, die Zählrichtung zu ändern. Dies bildet die Grundlage für das Erkennen von Zusammenhängen zwischen Addition und Subtraktion. Zählen wird als eine Strategie des Rechnens erkannt.

Das Erlernen der Zahlwortreihe ist ein langer Prozess, der sich über mehrere Jahre erstreckt (Fuson, 1988). Während sich die Niveaus 1 (String level) und 2 (Unbreakable list level) auf den Erwerb der ersten Zahlwörter und der Zahlwortreihe von ca. zwei- bis dreijährigen Kindern beziehen, können sich Kinder im Hinblick auf das Zählen in verschiedenen Zahlräumen auf unterschiedlichen Niveaus (3 bis 5) befinden. So können sie beispielsweise im Zahlenraum bis 20 vorwärts, rückwärts und in Schritten zählen, während ihnen dies im Zahlenraum bis 100 noch nicht gelingt. Nachdem die Kinder die unregelmässigen Zahlwörter zwischen eins und zwölf gelernt haben und die Zahlwortreihe bis zwanzig rezitieren können, können sie sich die anderen Zahlwörter aufgrund der dekadischen Regelmässigkeit mehr oder weniger erschliessen. Allerdings wirkt die Besonderheit der deutschen Zahlen-Syntax erschwerend auf den Erwerb der korrekten Zahlwortreihe. Auf diese Thematik wird in Abschnitt 4.4.2 näher eingegangen.

Mit dem Erwerb der Zahlwortreihe können Kinder nun auch die exakte Anzahl einer Menge bestimmen. Ob sich diese Zählfähigkeiten auf der Basis von im Kind angelegten Zählprinzipien entwickeln (Principles-before-Theorien) oder ob sie erlernt sind und sich durch Erfahrung verfestigen (Principles-after-Theorien), ist Gegenstand einer breiten Diskussion. Gelman und Gallistel (1986) gehen davon aus, dass den ersten Zählversuchen von Kindern konzeptuelle Kompetenzen zugrunde liegen. Sie postulieren als eine Art handlungsleitende Idee fünf Prinzipen, die aus ihrer Sicht angeboren sind und die Zählentwicklung massgeblich bestimmen. Diese Sichtweise ist heute überholt und wurde später auch von Gelman (1997) relativiert. Es wird davon ausgegangen, dass Kinder durch das aktive Zählen allmählich ihr Handlungsrepertoire erweitern und so zu einer sicheren Zählkompetenz gelangen, wobei sich nach und nach auch konzeptuelles Wissen entwickelt (Gasteiger, 2010; Moser Opitz, 2008). Die fünf Zählprinzipien von Gelman und Gallistel (1986) haben aber nach wie vor Gültigkeit und sind für das Bestimmen einer Anzahl von grosser Bedeutung. Die ersten drei Prinzipien werden als How-to-count-Prinzipien bezeichnet. Sie geben an, wie richtig gezählt wird und gelten als eigentliche Grundlage des Zählaktes (Gelman & Gallistel, 1986).

1. *One-one Principle* (Eindeutigkeitsprinzip)
Jedes Objekt wird beim Zählen nur einmal berücksichtigt. Das Kind muss die Menge in bereits gezählte und noch nicht gezählte Elemente aufteilen. Jedem zu zählenden Objekt muss dabei genau ein Zahlwort zugeordnet werden (Eins-zu-eins-Zuordnung).

2. *Stable-Order Principle* (Prinzip der stabilen Ordnung)
Dieses Prinzip meint, dass die beim Zählen benutzten Zahlwörter in einer stabilen, jederzeit wiederholbaren Reihenfolge verwendet werden. Das sichere Rezitieren der Zahlwortreihe ist Voraussetzung für das erfolgreiche Bestimmen einer Anzahl.

3. *Cardinal Principle* (Kardinalprinzip)
Das letzte Zahlwort, das beim Zählprozess genannt wird, gibt die Anzahl der jeweiligen Menge an. Das Konzept der Kardinalität ist dann vorhanden, wenn die Kinder erkannt haben, dass jede Zahl die bisher gezählten beinhaltet.

Die beiden weiteren Zählprinzipien beziehen sich auf das, was gezählt wird (permissions oder what-to-count principles). Sie sind dem eigentlichen Zählakt übergeordnet und ermöglichen Generalisierungen (Gelman & Gallistel, 1986).

4. *Abstraction Principle* (Abstraktionsprinzip)
Dieses Prinzip meint, dass die Art der Objekte für das Zählen nicht relevant ist. Das heisst, dass die ersten drei Zählprinzipien bei beliebigen Objekten angewandt werden können. Es können sowohl reale Gegenstände als auch Objekte ohne reale Existenz von homogenen oder heterogenen Gruppen gezählt werden.

5. *Order-Irrelevance Principle* (Prinzip der Irrelevanz der Anordnung)
Dieses letzte Prinzip besagt, dass die Reihenfolge, in der Objekte gezählt werden, keinen Einfluss auf das Ergebnis hat. Kinder müssen verstehen, dass ein zugeordnetes Zahlwort keine spezifische Eigenschaft des Elements darstellt und dass sich das Zählergebnis nicht ändert, wenn die Elemente in anderer Reihenfolge gezählt werden.

Wenn Kinder zählen und die Anzahl von Objekten genau bestimmen können, wenden sie diese Strategie auch für das Bestimmen der Summe von zwei oder mehreren Mengen an. Entwicklungsbedingt werden deshalb erste Additions- und Subtraktionsaufgaben zählend gelöst. Voraussetzung hierzu ist das Beherrschen der Zahlwortreihe, das sichere Zählen von Objekten und das Verständnis der Kardinalität (Moser Opitz, 2008). Dabei werden verschiedene Zählstrategien verwendet (z. B. Fuson, 1988; Gasteiger, 2010; Krajewski, 2008; Padberg & Benz, 2011):

Counting-all (vollständiges Auszählen):
Hierbei handelt es sich um die einfachste Strategie in der Bestimmung einer Summe mit Materialien oder mit Hilfe der Finger. Die Kinder bestimmen die Anzahl durch das Durchzählen aller Elemente. Manchmal bestimmen die Kinder zuerst die Anzahl beider Mengen separat und erst danach wird die Summe durch das Zählen aller Elemente bestimmt.

Für die Subtraktion werden zuerst die Elemente des Minuenden ausgezählt. Danach werden so viele Elemente abgezählt, wie der Subtrahend vorgibt. Das Ergebnis wird durch das Auszählen der übrigbleibenden Elemente bestimmt (separating from).

Counting-on from first (Weiterzählen vom ersten Summanden):
Die Elemente der ersten Menge (erster Summand) werden simultan erfasst und die Elemente der zweiten Menge (zweiter Summand) aufgezählt. Hierfür müssen die Kinder die Kompetenz zur Anwendung der flexiblen Zahlwortreihe besitzen.

Counting-on from larger (Weiterzählen vom grösseren Summanden):
Im Unterschied zur vorherigen Zählstrategie werden nicht mehr unbedingt die Elemente der ersten Menge als Ausgangswert genommen, sondern der grössere Summand. Die Ökonomie im Bestimmen einer Summe wird nochmals optimiert. Dieses Prinzip wird auch als *min* Strategie im Sinne von Zählschritte minimieren bezeichnet. Hier ist bereits ein implizites Wissen über das Vertauschen der Summanden vorhanden, welches erst viel später explizit als Kommutativgesetz kennengelernt wird.

Zur Lösung von Subtraktionsaufgaben wird ausgehend vom Minuenden um so viele Schritte rückwärts gezählt, wie der Subtrahend vorgibt (*counting down from*). Für die Strategie des Ergänzens (*counting up from given*) benötigen die Kinder bereits ein implizites Verständnis für die Umkehrbarkeit von Addition und Subtraktion. Sie zählen vom Subtrahenden so lange weiter, bis sie den Minuenden erreichen. Die Anzahl der Schritte liefert schliesslich das Ergebnis.

Auch wenn die genannten Zählstrategien innerhalb der mathematischen Kompetenzentwicklung Meilensteine darstellen, ist es wichtig, dass Kinder im Verlaufe des ersten Schuljahres allmählich weitere Rechenstrategien entdecken und nicht auf der Stufe des Zählens verweilen, da sonst die Gefahr der Verfestigung des zählenden Rechnens droht. Als weiterführende Strategien werden Recall oder Retrieval genannt (Gasteiger, 2010; Krajewski, 2008). Dabei wird die Summe bzw. die Differenz ohne zu zählen erfahrungsbedingt direkt bestimmt. Weitere Formen des Faktenabrufes umfassen die Strategien derived facts und decomposition (Gasteiger, 2010; Krajewski, 2008). Bei der erstgenannten Strategie wird das Ergebnis über eine andere, im Gedächtnis verfügbare Aufgabe abgeleitet. Mit der Strategie decomposition wird die Aufgabe in kleinere Teilaufgaben zerlegt, deren Lösungen abgerufen werden können.

In diesem Abschnitt wurde auf die approximative und die exakte Mengenrepräsentation eingegangen. Sie verkörpern zwei unterschiedliche Zugänge der Zahlbegriffsentwicklung. In mathematischen Kompetenzentwicklungsmodellen wird versucht, die Konzepte dieser beiden Mengenrepräsentationen systematisch darzustellen. Folgend werden solche Entwicklungsmodelle vorgestellt und miteinander verglichen.

4.2 Entwicklungsmodelle

Viele Entwicklungsmodelle mathematischer Kompetenzen beschränken sich auf den Beschreib einzelner Aspekte wie etwa das beschriebene Modell der Zählentwicklung (Fuson, 1988). Daneben gibt es neurokognitive Modelle wie beispielsweise das Triple-code-Modell von Dehaene und Cohen (1995). Dieses Modell wurde auf Basis von Erwachsenen mit einer Gehirnschädigung entwickelt und erklärt in erster Linie die Zahlenverarbeitung im Erwachsenenalter. Dabei wird zwischen drei unterschiedlichen Kodierungen von Zahl-Repräsentationen unterschieden: Zahlen als die numerische Grösse, als Zahlwörter und als Ziffer (Symbol). Basierend auf diesem Modell schlagen von Aster und Shalev (2007) ein Modell mit vier Stufen der numerischen Entwicklung dar, das sich auf die numerische Entwicklung im Kindesalter stützt. Die vier Stufen beinhalten basis-numerische Fertigkeiten, Zahlwortsymbole, visuell-arabische Zahlensymbole und Zahlenraumvorstellung. Dabei bezieht sich das Modell auf verschiedene Gehirnareale, die bei der Zahlenverarbeitung beteiligt sind und effizient zusammenarbeiten müssen. Neben diesen neuro-kognitiven Modellen liegen im deutschsprachigen Raum aktuell zwei kognitiv-entwicklungspsychologische Modelle vor, die die komplexe Wirklichkeit numerischer Entwicklung systematisch darzustellen und dabei verschiedene Aspekte zu integrieren versuchen: das Entwicklungsmodell der Zahl-Grössen-Verknüpfung nach Krajewski (2008) und das Modell der Entwicklung früher mathematischer Konzepte (Fritz & Ricken, 2008; Gerlach, Fritz, Ricken & Schmidt, 2007). Beide Modelle beschreiben den Erwerb des Wissens über Mengen, Zahlen und Rechenoperationen im Vorschulalter und sind für die vorliegende Arbeit zentral.

4.2.1 Entwicklungsmodell der Zahl-Grössen-Verknüpfung nach Krajewski (2008)

Ein erstes empirisch belastbares Entwicklungsmodell des Erwerbs früher mathematischer Kompetenzen wurde von Krajewski im Jahr 2003 generiert und in den darauffolgenden Jahren ständig weiterentwickelt (Schneider et al., 2016). Dieses Modell der Zahl-Grössen-Verknüpfung (ZGV) umspannt den Entwicklungsbereich ab der Geburt bis ins Grundschulalter.

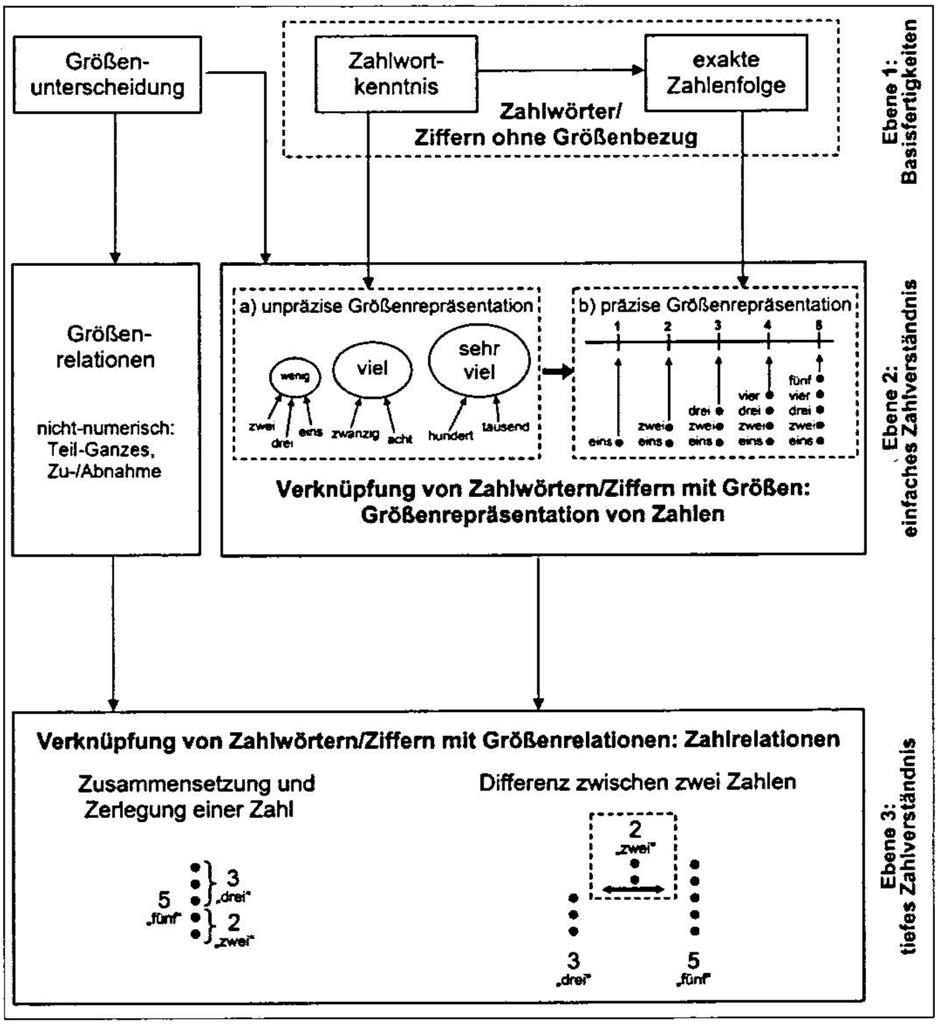

Abbildung 1: Entwicklungsmodell der Zahl-Grössen-Verknüpfung nach Schneider, Küspert und Krajewski (2016), S. 25

Die erste Ebene, *Basisfertigkeiten,* beschreibt einerseits die frühen Leistungen von Säuglingen zur Mengenunterscheidung (vgl. Abschnitt 4.1.1) und andererseits den Erwerb der Zahlwortreihe (vgl. Abschnitt 4.1.2). Diese beiden Lernprozesse finden unabhängig voneinander statt und sind noch nicht miteinander verknüpft. Die Mengenunterscheidung bezieht sich auf dieser Ebene noch nicht auf eine diskrete Anzahl, sondern auf den approximativen Vergleich der Quantität (mehr versus weniger) bzw. der Ausdehnung (viel versus wenig Raum). Sie ist vergleichbar mit dem protoquantitativen Schema des Vergleichens nach Resnick (1989). Mit Beginn des Sprechens lernen Kinder mit etwa zwei Jahren erste Zahlwörter kennen. Sie können auch schon exakte Zahlwortreihen von wenigen Zahlenfolgen wiedergeben, vergleichbar mit Level 1 (Ganzheitsauffassung der Zahlwortreihe) nach Fuson (1988). Sie wissen aber noch nicht, dass die exakte Menge durch Zählen ermittelt werden kann. Es steht primär der ordinale Zahlaspekt im Vordergrund. Die Verknüpfung von Zahlwörtern und Mengen erfolgt erst auf der zweiten Ebene.

Die zweite Ebene, *einfaches Zahlverständnis,* zeigt die Verknüpfung von Zahlwörtern mit Mengenrepräsentationen. Dies wird als bedeutender Meilenstein im Erwerb eines umfassenden Zahlbegriffes bezeichnet und läuft in zwei Phasen ab. Die Kinder erwerben im Alter von ca. drei Jahren zuerst ein unpräzises Anzahlkonzept, indem sie wissen, dass manche Zahlen für viel und andere für wenig stehen. So nutzen die Kinder beispielsweise die Zahl 1 für kleine Mengen und die Zahl 100 für grosse Mengen. In dieser Phase können sie aber noch keine exakten Unterscheidungen wahrnehmen und nicht bestimmen, ob beispielsweise 23 oder 25 weniger ist. Erst ein präzises Anzahlkonzept ermöglicht es, Elemente einzeln mit der Zahlenfolge zu verbinden und präzise anzugeben, dass beispielsweise 6 genau eins weniger ist als 7, aber eins mehr als 5. Jede einzelne Zahl korrespondiert exakt mit einer auszählbaren Menge (Schneider et al., 2016). Voraussetzung für das präzise Anzahlkonzept ist das Beherrschen der unflexiblen Zahlwortreihe (Level 2 nach Fuson, 1988) und die ersten drei Zählprinzipien nach Gelman und Gallistel (1978) (vgl. Abschnitt 4.1.2). Damit verfügen Kinder über das Kardinalitätsprinzip und haben verstanden, dass die letzte gezählte Zahl die Anzahl bzw. die Mächtigkeit einer Menge repräsentiert und alle vorhergehenden Zahlen beinhaltet. Unabhängig davon entwickelt sich auch das Verständnis für Mengen weiter. Die Kinder erkennen, dass Mengen durch die Zugabe bzw. Wegnahme von Objekten verändert werden können. Dies entspricht dem Schema des Vermehrens und Verminderns nach Resnick (1989). Zudem erkennen sie, dass Mengen in kleinere Mengen zerlegt und aus diesen auch wieder zusammengesetzt werden können. Dieses protoquantitative *Teil-Ganzes-Schema* (Resnick, 1989) gilt als Voraussetzung für die spätere Addition und Subtraktion.

Auf der dritten Ebene, *tiefes Zahlverständnis,* werden die Relationen zwischen Mengen schliesslich auch mit Zahlen beschreibbar. Das damit verbundene Verständnis für Beziehungen zwischen Zahlen stellt ein weiterer Meilenstein innerhalb der mathematischen Kompetenzentwicklung dar (Schneider et al., 2016). Im Alter von ca. sechs Jahren gelingt es den Kindern allmählich, Zahlen in kleinere Anzahlen zu zerlegen und wieder zusammenzusetzen. Die Kinder entwickeln auch ein Verständnis dafür, dass sich die Differenz zwischen zwei Zahlen durch eine dritte Zahl darstellen lässt. Diese Kompetenzen ermöglichen es, Zahlen in ihrer vollständigen Semantik zu verstehen und zum Rechnen zu verwenden.

4.2.2 Modell der mathematischen Kompetenzentwicklung nach Fritz und Ricken (2008)

Im Gegensatz zum Modell von Krajewski (2008) weist das Modell von Fritz und Ricken (2008) nicht nur drei, sondern fünf Niveaustufen auf und bezieht sich auf den sukzessiven Erwerb präziser arithmetischer Kompetenzen für den Altersbereich von vier bis acht Jahren (Fritz, Ehlert & Leutner, 2018).

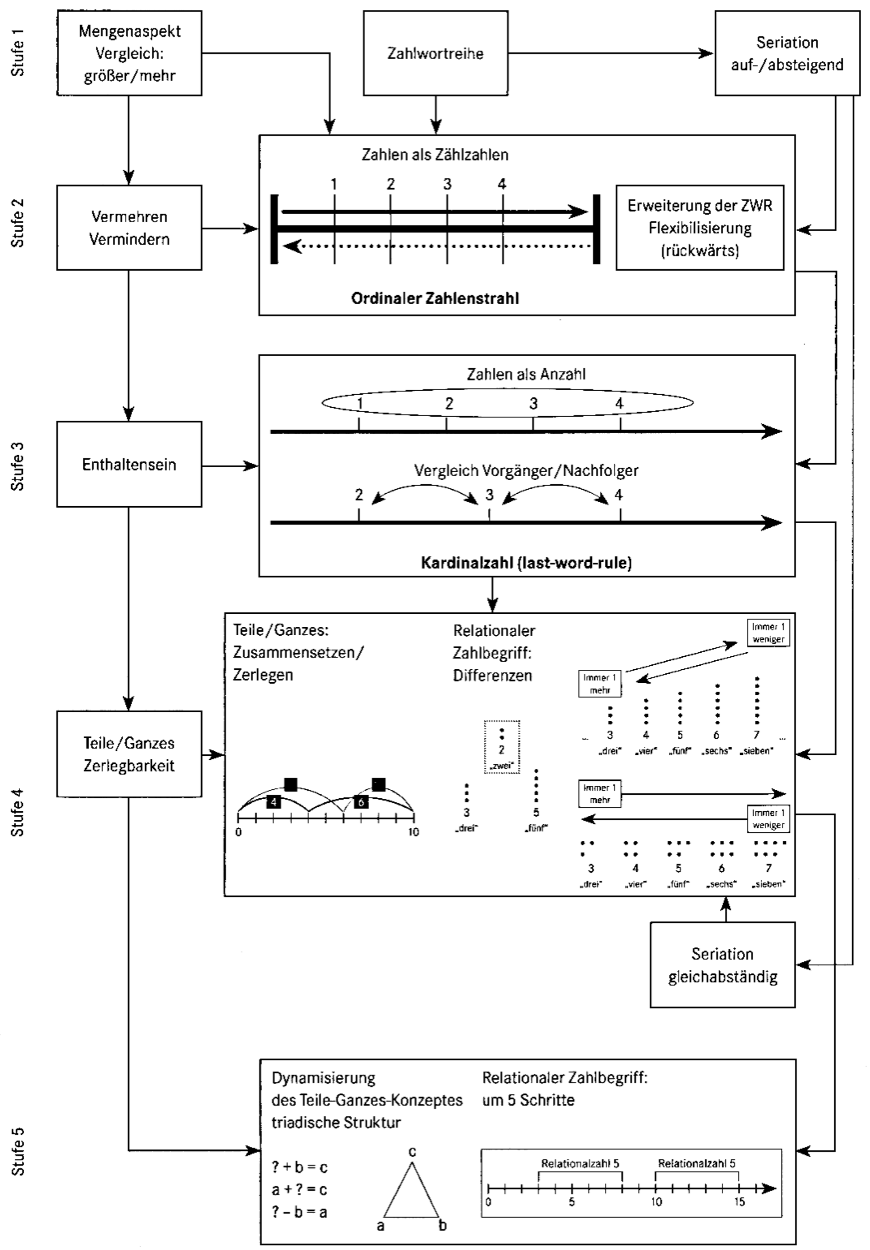

Abbildung 2: Kompetenzentwicklungsmodell nach Gerlach, Fritz, Ricken und Schmidt (2007), S. 15

Auf der ersten Stufe im Modell, *Niveau I: Zählzahl,* ist der Erwerb der Zahlwörter zentral. Es wird ein erstes Konzept präziser Zahlen erworben (Fritz et al., 2018). Dies stellt einen langen Prozess dar, wobei die Kinder anfänglich die Bedeutung der Zahlwörter noch nicht kennen. Sie lernen die Zahlwörter wie einen Reim aufzusagen, vergleichbar mit dem ersten Level (Ganzheitsauffassung der Zahlwortreihe) nach Fuson (1988). Allmählich erkennen sie die Zahlwörter als einzelne Einheiten und können Zählprozesse ausführen, wobei Zahlen- und Mengenkenntnisse noch unverbunden nebeneinander stehen. Eine weitere bedeutende

Kompetenz, die Kinder auf diesem Niveau erwerben, ist das Subitizing. Darunter wird das Erfassen kleiner Mengen bis ca. drei, vier Elementen auf einen Blick verstanden (vgl. Schema des Vergleichens von Resnick, 1989, Abschnitt 4.1.1). Auf dieser Stufe gelingt es den Kindern teilweise auch schon, den simultan erfassten Mengen ein Zahlwort zuzuordnen, aber ein inhaltliches Verständnis (Kardinalprinzip) ist noch nicht vorhanden. Auf diesem Niveau können Kinder ab einem Alter von zwei bis drei Jahren auch kleine Anzahlen der Grösse nach ordnen: Fähigkeit zur Seriation (Fritz & Ricken, 2008).

Auf der zweiten Stufe im Modell, *Niveau II: Ordinaler Zahlenstrahl,* wird das Wissen über Zahlen differenzierter. Es erfolgt eine Flexibilisierung der Zahlwortreihe, wobei auch das Rückwärtszählen und Weiterzählen von einer bestimmten Zahl möglich werden, vergleichbar mit dem Level 3 und 4 (teilweise flexible und flexible Zahlwortreihe) nach Fuson (1988) (vgl. Abschnitt 4.1.2). Die Kinder erwerben auf dieser Stufe einen mentalen Zahlenstrahl (Fritz et al., 2018). Die Zahlwörter stehen für eine bestimmte Position in der Reihe, die schrittweise grösser wird. Der Zahlenstrahl ist auf diesem Niveau noch nicht äquidistant, d. h. den Abständen zwischen den Zahlen wird noch keine Bedeutung zugemessen. Erst im Laufe der Schuleingangsphase wird der Zahlenstrahl in eine lineare und relationale Repräsentation überführt. Auf dieser Stufe ist es aber bereits möglich, Vorgänger und Nachfolger von Zahlen zu bestimmen, indem Zahlen über ihre Position in der Zahlwortreihe verglichen werden. Es gelingt nun, die grössere oder kleinere Zahl von zwei nahe beieinanderstehenden Zahlen (z. B. 5 oder 6) zu bestimmen und erste zählende Rechnungsaufgaben auszuführen. Auf dieser Stufe erkennen die Kinder auch, dass eine Menge durch Wegnehmen bzw. Hinzufügen verändert werden kann (vgl. Schema des Vermehrens und Verminderns nach Resnick, 1989, Abschnitt 4.1.1).

Beim Erreichen der dritten Stufe, *Niveau III: Kardinalität und Zerlegbarkeit,* besteht der wesentliche Schritt darin, dass die Anzahl der in der Menge enthaltenen Elemente als Mächtigkeit der Menge verstanden wird. Beim Zählen von Objekten steht beispielsweise die fünf nicht mehr nur für das fünfte Element, sondern für alle fünf gezählten Objekte. Voraussetzung für das exakte Bestimmen einer Anzahl ist das Einhalten der ersten drei Zählprinzipien nach Gelman und Gallistel (1978) (vgl. Abschnitt 4.1.2). Damit haben die Kinder das Kardinalkonzept erworben. Dieses Verständnis ermöglicht es, „dass Zahlen und Mengen nun auch über die Anzahl der Elemente ohne Zuhilfenahme der Eins-zu-Eins-Zuordnung miteinander verglichen werden können: 4 ist kleiner als 5, weil die Menge 4 weniger

Elemente enthält als die Menge 5“ (Fritz et al., 2018, S. 15f). Dieses erweiterte Zahlenverständnis schafft nun auch die Möglichkeit, dass Additionsaufgaben durch das Zuzählen der zweiten Menge zur ersten gelöst werden (counting-on-strategy, vgl. Abschnitt 4.1.2).

In der vierten Stufe im Modell, *Niveau IV: Enthaltensein und Klasseninklusion,* differenziert sich das Wissen über Mengen und deren Relationen weiter aus. Die Kinder erwerben als weiteren Meilenstein in der mathematischen Kompetenzentwicklung das Teil-Teil-Ganzes-Konzept, vergleichbar mit dem Teil-Ganzes-Schema nach Resnick (1989). Mengen repräsentieren eine Mächtigkeit, die sich aus einer Anzahl von Elementen zusammensetzt und damit in unterschiedliche Teilmengen zerlegt werden kann (Fritz et al., 2018). Dies gilt nicht nur für Mengen, sondern allmählich auch für Zahlen. Auch diese können unterschiedlich zerlegt und wieder zusammengesetzt werden. Die Differenz zwischen Zahlen kann nun ebenfalls bestimmt werden.

Mit dem Erreichen der fünften Stufe im Modell, *Niveau V: Relationalität,* können Beziehungen zwischen den Konzepten der Ordinaliät (Niveau II) und der Kardinalität (Niveau III) hergestellt werden. „Die Zahlwortreihe wird als eine Sequenz aufeinanderfolgender Zahlen verstanden, in der jede Zahl eine Vorgänger- und Nachfolgerzahl hat, die um 1 kleiner bzw. um 1 grösser ist. Die Abstände zwischen zwei aufeinander folgenden Zahlen sind somit immer gleich“ (Fritz et al., 2018, S. 17). Mit diesem Wissen können Zahlen nun auch für kongruente Intervalle stehen und einen Abschnitt auf dem Zahlenstrahl bzw. eine Beziehung zwischen zwei Mengen darstellen. Zusätzlich gelingt es Kindern auf dieser Stufe, von einer beliebigen Zahl um eine Anzahl Schritte weiter oder zurück zu zählen, vergleichbar mit der flexiblen bzw. vollständig reversiblen Zahlwortreihe nach Fuson (1988) (vgl. Abschnitt 4.1.2).

Neu erwähnen Fritz et al. (2018) eine weitere Stufe, *Niveau VI: Zahlen gleichmässig bündeln*, die im obigen Modell noch nicht berücksichtigt wurde. Kinder auf dieser Stufe sind in der Lage, unterschiedliche gleichmächtige Zerlegungen einer Zahl zu finden, und können umgekehrt Zahlen als Bündeleinheiten verstehen. Dieses Wissen ist insbesondere für das Verständnis der Multiplikation (z. B. entspricht 3 x 5 drei Fünferbündeln), der Division (z. B. bedeutet 12 : 4, dass die 12 in Viererbündel zerlegt werden kann) und des Stellenwertsystems (z. B. die 4 von 42 bedeutet vier Zehnerbündel) wesentlich.

4.2.3 Diskussion der beiden Entwicklungsmodelle

Zwischen dem Entwicklungsmodell der Zahl-Grössen-Verknüpfung (ZGV) nach Krajewski (2008) und dem Modell der mathematischen Kompetenzentwicklung von Fritz und Ricken (2008) sind deutliche Parallelen erkennbar. Beide Modelle gehen von der Ausdifferenzierung spezifisch mathematischer Kompetenzen aus, die in der frühen Kindheit beginnt. Dabei setzt das ZGV-Modell (Krajewski, 2008) den Schwerpunkt auf die Entwicklung des Zahlen- und Mengenverständnisses, während das Modell von Fritz und Ricken (2008) die numerische und arithmetische Entwicklung fokussiert (Hildenbrand, 2016). Beide Modelle integrieren aber die approximative und exakte Mengenrepräsentation. Entsprechend sind in beiden Modellen Bezüge zu den protoquantitativen Schemata von Resnick (1989), den Stufen der Zählentwicklung nach Fuson (1988) und den Zählprinzipien nach Gelman und Gallistel (1978) ersichtlich. Insgesamt entsprechen die beiden ersten Stufen bei Fritz und Ricken (2008) der ersten Ebene Krajewskis (2008), wobei das Vermehren und Vermindern von Mengen bei Krajewski erst auf der zweiten Ebene angesiedelt ist. Die Stufen 3 und 4 von Fritz und Ricken (2008) lassen sich mit der zweiten Ebene Krajewskis (2008) und die fünfte Stufe mit der dritten Ebene vergleichen. Trotz der Beschreibung von Ebenen und Stufen verstehen sich beide Modelle nicht streng hierarchisch. Da der Umgang mit Zahlenfolgen in kleineren Zahlenräumen schon elaborierter ist als in grösseren, kann sich ein Kind in unterschiedlichen Zahlenräumen in verschiedenen Ebenen bzw. auf verschiedenen Stufen befinden (Schneider et al., 2016). Ebenso begreifen beide Modelle den Erwerb des Zahlbegriffes als einen aktiven Prozess, der sich über den Umgang mit Zahlen und insbesondere über das Zählen entwickelt. So postulieren beide Modelle das Beherrschen der Zahlwortfolge und das Verstehen des Kardinalprinzips als grundlegende Voraussetzungen für das spätere Addieren und Subtrahieren. Das Entwicklungsmodell der Zahl-Grössen-Verknüpfung (Krajewski, 2008) wurde bisher nicht empirisch überprüft (Fritz et al., 2018). Es bildet aber die Grundlage eines mehrfach positiv evaluierten Förderprogramms *Mengen, Zählen, Zahlen (MZZ)* (Ennemoser, Sinner & Krajewski, 2015). Das Modell von Fritz und Ricken (2008) hingegen wurde empirisch validiert. In zwei Längsschnittstudien konnten die im Modell beschriebenen Entwicklungsverläufe empirisch dargestellt werden (Fritz et al., 2018).

Die in den beiden Modellen beschriebene Entwicklung mathematischer Kompetenzen beginnt mit der Geburt. Entsprechend verfügen Kinder bereits über ein breites Spektrum an mathematischen Kompetenzen, wenn sie in den Kindergarten kommen. Diese werden im Verlauf der Kindergartenzeit weiterentwickelt und vertieft. Im nächsten Kapitel werden die mathematischen Kompetenzen, über die Kinder im Kindergarten verfügen sollen, detailliert dargestellt und mögliche Stolpersteine beim Erwerb dieser Kompetenzen aufgezeigt.

4.3 Erkenntnisse für den Kindergarten: Kompetenzen und Stolpersteine

Wullschleger (2017) nennt sieben Kompetenzbereiche, die Kinder im Kindergarten erwerben sollen bzw. die während der Kindergartenzeit aufzubauen sind. Empirische Befunde zeigen, dass der Erwerb dieser mathematischen Kompetenzen mit Stolpersteinen verbunden sein kann. Folgend werden die Kompetenzen und möglichen Schwierigkeiten beim Erwerb beschrieben.

4.3.1 Mengen vergleichen

Wie in den Entwicklungsmodellen aufgezeigt, können Kinder von frühster Kindheit an Unterschiede zwischen Mengen erkennen (vgl. Ebene 1 bei Krajewski (2008) und Stufe 1 bei Fritz und Ricken (2008)). Während die Differenzierung anfänglich nur bei kleinen Mengen oder zwischen Mengen mit deutlich unterschiedlichen Anzahlen gelingt, differenziert sich diese Kompetenz im Verlaufe des Kindergartens weiter. Es gibt unterschiedliche Vorgehensweisen, wie Kinder Mengen vergleichen können. In der Regel finden sie selbst Strategien. Wenn Kinder aber nicht von sich aus Vorgehensweisen zum Vergleichen entdecken und anwenden, müssen sie dabei angeleitet und die Vorgehensweisen geübt werden. Eine erste Strategie ist der approximative Grössenvergleich (vgl. Abschnitt 4.1.1). Mengen können durch Schätzen verglichen werden, wenn die Mengen nur wenige Elemente enthalten, die Differenz der Anzahl Elemente zwischen den Mengen gross ist oder die Elemente beider Mengen in der gleichen Struktur dargestellt werden. Beispielsweise lässt sich im Zwanzigerfeld schnell erkennen, dass 17 Punkte mehr sind als 15, wenn die Punkte hingegen ohne Struktur dargestellt sind, wäre ein approximativer Mengenvergleich kaum möglich. Eine weitere Strategie ist das Vergleichen von Mengen mittels Eins-zu-eins-Zuordnung. Dabei wird jedem Element der einen Menge genau ein Objekt der jeweils anderen Menge zugeordnet (Benz et al., 2015). Wenn reale Objekte zweier Mengen verglichen werden sollen, beispielsweise rote und blaue Murmeln, kann die Eins-zu-eins-Zuordnung handelnd vollzogen werden, indem zu jeder roten eine blaue Murmel gelegt wird. Wenn Mengen auf ikonischer Ebene verglichen werden sollen, kann die Eins-zu-eins-Zuordnung visuell vorgenommen werden. Dabei dürfen die zu vergleichenden Mengen jedoch nicht zu viele Elemente enthalten. Es ist darüber hinaus auch möglich, ein Element der einen Menge mit einem Element der anderen Menge durch einen Strich zu verbinden. Wenn die zu vergleichenden Mengen weder durch Schätzen noch durch eine Eins-zu-eins-Zuordnung verglichen werden können, müssen die Elemente beider Mengen durch Zählen exakt bestimmt und erkannt werden, welche der beiden Zahlen die grössere ist. Die mit der Anzahlbestimmung verbundenen Stolpersteine werden in Abschnitt 4.4.4 beschrieben. Beim Vergleichen von Mengen sollen die Kinder im Kindergarten die Konzepte grösser/kleiner, mehr/weniger, gleich viel, am meisten und am wenigsten kennen lernen (Wullschleger, 2017).

4.3.2 Erwerb der Zahlwortreihe

Kinder lernen nicht erst im Kindergarten das Rezitieren der Zahlwortreihe (vgl. Entwicklung der Zahlwortreihe nach Fuson (1988), Abschnitt 4.1.2). Allerdings gibt es Kinder, die damit noch wenig vertraut sind. Deshalb sind verschiedene Zählaktivitäten im Kindergarten von zentraler Bedeutung. Manche Kinder erlernen das Aufsagen der Zahlwortreihe in Verbindung mit einer Zählhandlung, andere bevorzugen das Üben der Zahlwortreihe auf rein verbaler Ebene (Scherer & Moser Opitz, 2010). Beim Schuleintritt sollten die Kinder die Zahlwortreihe bis 20 sicher beherrschen, von 10 rückwärts zählen können und auch von einer beliebigen Zahl weiter zählen können (vgl. Abschnitt 2.1.2). Das fehlerfreie Rezitieren der Zahlwortreihe ist mit einer hohen Gedächtnisleistung verbunden und stellt generell eine grosse Anforderung dar. Die Bildungsregel der deutschen Zählwörter erschwert das korrekte Zählen zusätzlich (Schneider et al., 2016). Die unregelmässigen Zahlwörter bis 12 müssen gelernt werden. Danach kann von 13 bis 19 nach einem einheitlichen Prinzip gezählt werden. Die Zehnerzahlen sind verwandt mit den Zahlen des ersten Zehners, müssen jedoch mit der Endung -zig versehen und somit auch erlernt werden. Wobei auch hier Unregelmässigkeiten vorkommen: zwanzig anstelle von zweizig (Scherer & Moser Opitz, 2010). Zusätzlich erschweren auch Unregelmässigkeiten in der Aussprache der deutschen Zahlwörter den Erwerb der Zahlwortreihe. „Ab 13 werden zuerst die Einer, dann die Zehner genannt, ab 21 zusätzlich durch die Silbe ›und‹ verbunden“ (Scherer & Moser Opitz, 2010, S. 106). Solche Unregelmässigkeiten erhöhen das Fehlerrisiko beim Zählen. In der deutschen Syntax existieren aber auch logisch nachvollziehbare, besondere Formen der Auslassung wie z. B. die sogenannten Schnapszahlen (22, 33, 44 usw.). Wenn die Kinder die Zahl vor der Schnapszahl aussprechen (z. B. zwei-und-dreissig) finden sich die beiden Zahlen 2 und 3 in der bekannten Reihenfolge. Dies lässt die Kinder vermuten, dass als nächste Zahl eine 4 kommt. Folglich lassen sie die Zahl 33 aus und nennen stattdessen die 34 (Scherer & Moser Opitz, 2010). Weitere Untersuchungen haben gezeigt, dass der Übergang über den Zehner oft schwierig zu meistern ist (Moser Opitz, 1999; Selter & Spiegel, 2007). Anstelle von neunundzwanzig wird beispielsweise das Zahlwort neunzig genannt. Manchmal erfinden die Kinder auch neue (logische) Zahlwörter wie einzehn, zweizehn, elfizig oder lassen eine Zahl aus, weil sie dieses Zahlwort noch nicht kennen oder vergessen haben. Diese Ausführungen verdeutlichen die Hürden beim Erwerb der korrekten Zahlwortreihe. Wenn Kinder eine andere Erstsprache als Deutsch haben, wird dieser Erwerb noch zusätzlich erschwert. So weist die deutsche Zahlwortreihe grössere Unregelmässigkeiten auf als beispielsweise die türkische oder asiatische (Park, 2000). In diesem Zusammenhang konnten Moser Opitz, Ruggiero und Wüest (2010) nachweisen, dass Kinder mit einer anderen Erstsprache als Deutsch geringere Zählkompetenzen aufwiesen als deutschsprachige Kinder. Sie evaluierten die Zählfähigkeit von insgesamt 355 Kindergartenkindern in der Deutschschweiz. Dabei verwendeten Kinder mit einer anderen Erstsprache als Deutsch häufiger falsche Zahlwörter als deutschsprachige Kinder. Entgegen den Erwartungen wurde ersichtlich, dass Kinder mit einer anderen Erstsprache als Deutsch die Zahlwortreihe auf Deutsch besser beherrschten als in ihrer eigenen Erstsprache. Eine Ausnahme bildeten türkische Kinder, die in beiden Sprachen etwa identische Zählfertigkeiten aufwiesen.

4.3.3 Zahlenreihenfolge

Wenn Kinder sich mit der Zahlenreihenfolge auseinandersetzen, steht der ordinale Zahlaspekt im Vordergrund. Dabei nimmt jede Zahl innerhalb einer festen Reihenfolge eine klar definierte Position ein (Hasemann & Gasteiger, 2014). Gleichwohl kann es für Kinder schwierig sein, Zahlen der Grösse nach zu ordnen, einzelne Zahlen in die Zahlenreihenfolge einzufügen oder zu einer vorgegebenen Zahl die Nachbarzahlen zu bestimmen. Um diese Kompetenzen zu erwerben, müssen Kinder immer wieder die Gelegenheit erhalten, sich mit der Zahlenreihenfolge sowohl verbal als auch visuell auseinanderzusetzen. Dabei helfen Zähllieder, das Aufsagen der Zahlwortreihe vorwärts, rückwärts und in Schritten (Moser Opitz, 2008) oder Spiele, bei denen die Zahlen in die richtige Reihenfolge gelegt werden müssen, wie beispielsweise das *Fünferraus* (vgl. Abschnitt 7.2). Neben der Zählzahl gehört auch die Ordnungszahl zum Ordinalzahlaspekt (Benz et al., 2015). Hier wird die Position eines Elementes in einer festen Reihenfolge bestimmt, beispielsweise „das vierte Kind in der Reihe“ oder „Mia wurde zweite“. Der Aspekt der Ordnungszahl fällt den Kindern in der Regel leichter als Vorgänger oder Nachfolger von Zahlen zu bestimmen, da er alltagsnäher ist.

4.3.4 Anzahlbestimmung

Beim Bestimmen der Anzahl einer Menge steht der Kardinalzahlaspekt im Vordergrund (Hasemann & Gasteiger, 2014). Kinder können die Anzahl einer Menge auf unterschiedliche Weise bestimmen. Häufig zählen sie die einzelnen Elemente der Menge beispielsweise durch Antippen oder Verschieben. Um eine Anzahl exakt durch Zählen zu bestimmen, müssen die Prinzipien nach Gelman und Gallistel (1986) eingehalten werden (vgl. Abschnitt 4.1.2). Dabei gilt es einige Stolpersteine zu überwinden. Es gibt Kinder, die asynchron zählen. Die Eins-zu-eins-Zuordnung von Zahlwort und Objekt gelingt dann nicht und das Eindeutigkeitsprinzip ist verletzt (vgl. Abschnitt 4.1.2). Ein Zahlwort wird genannt, aber keinem Objekt zugeordnet, oder zwei Objekten wird nur ein Zahlwort zugeordnet. Letztes ist oft beim Zahlwort sieben der Fall, weil es innerhalb des ersten Zehners das einzige zweisilbige Zahlwort ist (Benz et al., 2015). Die Ursachen für diese Zählfehler liegen in der ungenügenden Koordination zwischen Zeigen und Objekt (Auge-Hand-Koordination) oder in der ungenügenden Koordination zwischen Objekte antippen und Zahlwörter zuordnen (Hess, 2012). Wenn die zu zählenden Elemente ungeordnet sind, kann es auch vorkommen, dass Kinder einige Elemente mehrmals zählen. In diesem Fall ist es sinnvoll, die Objekte in eine Reihe zu legen und jedes gezählte Objekt ein Stück zu verschieben. Manchmal ist auch zu beobachten, dass Kinder die Objekte zwar richtig zählen, aber auf die Frage „Wie viele sind es?“ wieder mit Zählen beginnen. In diesem Fall haben sie das Kardinalprinzip noch nicht verstanden und wissen nicht, dass das letzte Zahlwort die Anzahl der Objekte repräsentiert (vgl. Abschnitt 4.1.2). Allerdings kann es auch sein, dass Kinder auf die Frage „Wie viele?“ richtig antworten, aber die Bedeutung der Kardinalität noch nicht wirklich erkannt wird (Hess, 2012). Das Konzept der Kardinalität ist erst vorhanden, wenn die Kinder erkannt haben, dass jede Zahl die bisher gezählten beinhaltet. Dies spiegelt sich auch darin wider, dass Zahlen in Beziehung zu anderen Mengen gesetzt werden und ein Bewusstsein dafür vorhanden ist, dass beispielsweise 8 eins mehr ist als 7. Das Zählen von Gegenständen ist zudem sehr fehleranfällig. Auch Kinder, die eigentlich schon über eine sichere Zählkompetenz verfügen, können sich schnell einmal verzählen. Neben dem Bestimmen der Anzahl durch Zählen kann die Anzahl von Objekten auch durch simultane bzw. quasi-simultane Erfassung bestimmt werden. Bis zu einer Anzahl von drei, vier Elementen können Kinder die Anzahl auf einen Blick bestimmen. Diese Fähigkeit wird auch Subitizing genannt (vgl. Abschnitt 4.1.1). Wenn grössere Anzahlen auf einen Blick erkannt werden sollen, müssen die einzelnen Objekte zunächst in kleinere Gruppen zusammengefasst werden. Beispielsweise können sechs Objekte als zwei Dreiergruppen wahrgenommen werden (quasi-simultane Erfassung, vgl. Abschnitt 4.1.1). Grössere Anzahlen können auch auf einen Blick wahrgenommen werden, wenn die Elemente in einer bekannten Struktur dargestellt sind (strukturierte Anzahlerfassung). Um die Anzahl der Punkte beispielsweise auf dem Würfel oder auf dem Zehner- und Zwanzigerfeld auf einen Blick zu erkennen, müssen die Kinder vorgängig genügend Zeit haben, die Punkte zu zählen und die Struktur zu erkennen (Moser Opitz, 2012).

Neben dem in Abschnitt 4.4.3 erwähnten ordinalen Zahlaspekt und dem hier angesprochenen kardinalen Zahlaspekt werden in der Literatur (Hasemann & Gasteiger, 2014; Padberg & Benz, 2011; Radatz & Schipper, 1983) noch vier weitere Zahlaspekte genannt:

- Der *Masszahlaspekt* gibt die Grösse von Zahlen im Zusammenhang mit einer dazugehörigen Masseinheit an (z. B. fünf Minuten). Kinder können hier beispielsweise die Erfahrung machen, dass fünf Kilogramm Äpfel etwas anderes bedeutet als fünf Äpfel.
- Der *Operatoraspekt* beschreibt die Anzahl der Wiederholungen von Vorgängen oder Handlungen (z. B. die Kirchenglocke schlägt zweimal) und liefert die Antwort auf die Frage „Wie oft?“.
- Der *Rechenzahlaspekt* meint verschriftliche Zahlen, die als Symbole unter Einhaltung gewisser Regeln zum Rechnen benutzt werden (z. B. 8 + 5 = 13).
- Der *Codierungsaspekt* bezieht sich auf Zahlen, die eine Labelfunktion besitzen. Darunter werden beispielsweise Zahlen und Ziffernfolgen in Telefonnummern, Autokennzeichen oder Postleitzahlen verstanden.

Kindergartenkinder müssen nicht bewusst zwischen den verschiedenen Zahlaspekten unterscheiden. Sie sollen aber die Erfahrung machen, dass Zahlen im Alltag in verschiedenen Situationen gebraucht werden (z. B. im Zusammenhang mit Massen) und an verschiedenen Orten zu finden sind (z. B. auf Autonummernschild oder Hausnummern).

4.3.5 Mengen-Zahlen-Zuordnung

Obwohl Kinder schon lange bevor sie den Kindergarten besuchen in ihrem Alltag mit Zahlen in Berührung kommen, wird das Thematisieren von Zahlen im Kindergarten immer wieder kritisiert (Wullschleger, 2017). Aus der Forschung ist aber bekannt, dass die Zahlen-Mengen-Verbindung für die mathematische Entwicklung von grosser Bedeutung ist (Clarke, Clarke, Grüßing & Peter-Koop, 2013). Bruner (1971) plädiert zudem dafür, dass Lerninhalte zu

Mengen und Zahlen auf drei verschiedenen Repräsentationsebenen miteinander in Beziehung zu setzen sind, damit wirksames und ganzheitliches mathematisches Lernen möglich ist. Das sogenannte EIS-Prinzip (Bruner, 1971) fordert, dass Erfahrungen auf der enaktiven (handelnden), der ikonischen (bildhaften) und der symbolischen Ebene gemacht werden können. Nach Benz et al. (2015) ist es notwendig, dass mathematische Inhalte von einer Ebene in die andere übersetzt werden, damit ein Verständnis für mathematische Inhalte aufgebaut werden kann. Diese Übersetzung wird als intermodaler Transfer bezeichnet. Entsprechend sind im Kindergarten Lerngelegenheiten erforderlich, in denen Handlungen (beispielsweise fünf Murmeln in einen Reifen legen), mit einem Bild (fünf Murmeln in einem Reifen zeichnen) und mit einem Symbol (5 Murmeln) verbunden werden können. Der Transfer zwischen Handlung und Symbol bzw. zwischen Bild und Symbol ist nicht für alle Kinder leicht nachzuvollziehen. Übersetzungen von einer Darstellungsform in eine andere helfen dabei, ein Verständnis für Symbole aufzubauen, und sollten im Kindergarten immer wieder bei für die Kinder bedeutsamen Situationen thematisiert werden (Benz et al. 2015).

4.3.6 Zerlegen und Zusammensetzen von Mengen

Auf die Bedeutung des Zerlegens und Zusammensetzens von Mengen wurde bereits im Zusammenhang mit den Entwicklungsmodellen (vgl. Abschnitt 4.2) und dem Teil-Ganzes-Schema von Resnick (1989) in Abschnitt 4.1.1 aufmerksam gemacht. Das Zerlegen einer Menge in Teilmengen bzw. das Zusammensetzen von Mengen zu einer Gesamtmenge ermöglicht ein vertieftes Verständnis für Beziehungen zwischen Mengen und Teilmengen. Dieses lässt sich in einem nächsten Schritt auch auf Zahlen übertragen. Kinder gelangen so zu Erkenntnissen über Beziehungen zwischen Zahlen. Sie erkennen, dass man die Zahl 5 in 2 und 3 oder in 4 und 1 zerlegen kann und dass man aus 2 und 3 wieder 5 erhält (Benz et al. 2015). Auch dieses Wissen erwerben nicht alle Kinder ohne spezifische Anregungen. Im Kindergarten sollen deshalb immer wieder Lerngelegenheiten angeboten werden, in denen gemeinsam über Zahlen und Beziehungen zwischen Zahlen nachgedacht und gesprochen wird.

4.3.7 Erstes Rechnen

Wenn Kinder die genaue Anzahl bestimmen können, wenden sie diese Strategie auch für das Bestimmen der Summe von zwei oder mehreren Mengen an. Diese ersten Additionsaufgaben werden zählend gelöst. Im Kindergarten ist dies oft der Fall, wenn ein Würfelspiel mit zwei

Würfeln gespielt wird und die Summe beider Würfelaugen bestimmt werden soll. Anfänglich zählen die Kinder die Augen beider Würfel durch (vgl. Abschnitt 4.1.2). Wenn Kinder in der Strategie *Counting-all* verhaften bleiben, können Inputs von Fachkräften der folgenden Art hilfreich sein: „Gibt es eine schnellere Möglichkeit die Punkte beider Würfel zu bestimmen?“ oder „Beim ersten Würfel kannst du die Anzahl auf einen Blick erkennen, hier musst du nicht mehr zählen“. Damit können Kinder die weiteren Strategien *Counting-on from first* und *Counting-on from larger* erwerben. Mit der Zeit gelingt es einigen Kinder auch, die Summe beider Würfel zu bestimmen, ohne zu zählen. Viele Kinder können vor Schuleintritt auch schon einfache Rechnungen auf der symbolischen Ebene lösen, beispielsweise wenn mit zwei Zahlenwürfeln anstelle von Augenwürfeln gespielt wird (Hasemann & Gasteiger, 2014).

4.4 Zusammenfassung

In diesem Kapitel wurde aufgezeigt, wie sich mathematische Kompetenzen von der Geburt bis zum Kindergartenalter und leicht darüber hinaus entwickeln. Dabei wurde zwischen der approximativen und der exakten Mengenrepräsentation unterschieden. In der approximativen Mengenrepräsentation wurde ein Bezug zu Habituationsexperimenten mit Säuglingen hergestellt. In diesen Experimenten gab es Hinweise darauf, dass bereits Säuglinge Mengen wahrnehmen und unterscheiden können, wenn sich diese in der Quantität deutlich voneinander abheben. Im Laufe des ersten Lebensjahres entwickelt sich diese Kompetenz der Mengenunterscheidung weiter. Kinder sind zunehmend in der Lage, zwischen immer grösseren Mengen mit immer kleineren Differenzen zu unterscheiden. Neben diesem Schema des Vergleichens lernen Kinder im Verlaufe der ersten Lebensjahre auch die Konzepte des Vermehrens und Verminderns bzw. des Teil-Ganzes kennen. Sie erkennen, dass sich Mengen verkleinern bzw. vergrössern und in Teilmengen zerlegen lassen. Diese protoquantitativen Schemata differenzieren sich mit dem Erwerb der Sprache weiter aus und sind für den Erwerb eines tragfähigen Zahlbegriffs von Bedeutung. Im Zusammenhang mit der exakten Mengenrepräsentation wurde auf die Bedeutung des Lernens der Zahlwortreihe und der Zählkompetenz aufmerksam gemacht. Die Kinder erlernen die exakte Zahlwortreihe über verschiedene Entwicklungsstufen und über mehrere Jahre. Diese Kompetenz bildet eine zentrale Voraussetzung zur erfolgreichen Anzahlbestimmung. Daneben müssen die Kinder das Prinzip der Eindeutigkeit (Eins-zu-eins-Zuordnung: jedem Gegenstand ein Zahlwort) und der Kardinalität (das letztgenannte Zahlwort repräsentiert die Anzahl) verstanden haben. Mit dem Erwerb der Zählkompetenz können Additions- und Subtraktionsaufgaben zählend gelöst werden. Dabei wenden Kinder verschiedene Zählstrategien an, die das Bestimmen einer Summe bzw. einer Differenz immer weiter optimieren.

Die approximative und die exakte Mengenrepräsentation verkörpern zwei unterschiedliche Zugänge zur Zahlbegriffsentwicklung. In mathematischen Kompetenzentwicklungsmodellen wird versucht, die Konzepte dieser beiden Mengenrepräsentationen systematisch darzustellen. Die beiden im deutschsprachigen Raum für den Vorschulbereich fundierten Modelle von Krajewski (2008) sowie Fritz und Ricken (2008) sind sich im Wesentlichen ähnlich. Beide stellen den Erwerb des Wissens über Mengen, Zahlen und erste Rechenoperationen in qualitativ unterschiedlichen Kompetenzstufen bzw. Kompetenzebenen dar, wobei diese nicht streng hierarchisch zu verstehen sind. Ein Kind kann je nach Zahlenraum zeitgleich auf unterschiedlichen Ebenen oder Stufen der Kompetenzentwicklung stehen. Während der Kindergartenzeit sollen Kinder die in den Entwicklungsmodellen beschriebenen Kompetenzen erwerben. Diese bilden die Grundlage für weitere mathematische Tätigkeiten in der Schule. Dazu gehören das Vergleichen von Mengen, der Erwerb der Zahlwortreihe, Kenntnisse der Zahlenreihenfolge, das korrekte Bestimmen einer Anzahl, das Zuordnen von Mengen und Zahlen, das Zerlegen und Zusammensetzen von Mengen und erstes Rechnen. Mit dem Erwerb dieser Kompetenzen können auch Stolpersteine verbunden sein. Kinder begehen beispielsweise Zählfehler, haben keine Strategie, wie sie zwei Mengen vergleichen können oder zeigen Schwierigkeiten beim Bestimmen einer Anzahl, weil sie asynchron zählen. Um solche Schwierigkeiten überwinden zu können, ist es von Relevanz, dass im Kindergarten ein breites mathematisches Lernangebot besteht und Kinder auf vielfältige Art und Weise verschiedene mathematische Tätigkeiten immer wieder anwenden und üben können.

Trotzdem verläuft die mathematische Entwicklung im Kindergarten nicht bei allen Kindern gleich und nicht alle Kinder verfügen am Ende der Kindergartenzeit über die gleichen Kompetenzen. Individuelle und kontextuelle Faktoren können die Entwicklung beeinflussen. Im nächsten Kapitel wird der aktuelle Forschungsstand zu diesen Zusammenhängen aufgearbeitet.

5 Zusammenhänge mathematischer Kompetenzen mit Kontextfaktoren

Im vorherigen Kapitel wurde beschrieben, wie sich mathematische Kompetenzen entwickeln. Diese Entwicklung verläuft aber nicht bei allen Kindern gleich. So zeigen verschiedene Studien, dass im Kindergarten und in der ersten Klasse die Unterschiede zwischen Kindern in Bezug auf ihre mathematischen Kompetenzen sehr gross sind (z. B. Hasemann & Gasteiger, 2014; Schmidt, 1982; Stamm, 2005). Krajewski und Schneider (2006) wiesen nach, dass die Hälfte der von ihnen untersuchten Kindergartenkinder sechs Monate vor dem Schuleintritt bis 20 zählen konnten, 56 % der Kinder Aufgaben zur numerischen Seriation und 63 % der Kinder Aufgaben zum Mengenvergleich richtig lösen konnten. Dies bedeutet aber gleichzeitig auch, dass viele Kinder diese Aufgaben noch nicht bearbeiten konnten und die Heterogenität entsprechend gross war. In der Untersuchung von Moser, Stamm und Hollenweger (2004) mit 2000 Kindern aus dem Raum Zürich kannten mehr als die Hälfte der Kinder zu Beginn der ersten Klasse die Zahlen bis 20, konnten eine unvollständige Zahlenreihe bis 20 mit den entsprechenden Zahlen ergänzen und einfache Additionsaufgaben bis 10 lösen. Daneben gab es aber auch Kinder, die noch keine oder nur wenige Zahlen im Zahlenraum bis 10 kannten und Anzahlen bis 20 nicht durch Zählen bestimmen konnten. Gleichzeitig waren auch Kinder darunter, die die Zahlen bis 100 kannten und im Zahlenraum bis 20 und grösser subtrahierten und addierten. Ähnliche Ergebnisse liegen auch aus Untersuchungen in Deutschland vor. Bei der Erprobung des Osnabrücker Tests zur Zahlbegriffsentwicklung mit mehr als 300 Kindern konnten unmittelbar vor Schulbeginn 58 % der Kinder 20 Klötze richtig abzählen (Hasemann & Gasteiger, 2014), was gleichzeitig bedeutet, dass 42 % der Kinder diese Aufgaben kurz vor Schulbeginn noch nicht lösen konnten. In der Literatur werden unterschiedliche Erklärungen für diese Heterogenität bzw. diese verschiedenen Entwicklungsverläufe diskutiert. Zentral scheint dahingehend der Einfluss von Kontextfaktoren zu sein. Unter Kontextfaktoren werden in Anlehnung an die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) sowohl individuelle, personenbezogene Faktoren als auch Umweltfaktoren verstanden (Schuntermann, 2001). Die Forschung zu den Zusammenhängen mathematischer Kompetenzen mit solchen Faktoren wurde in den letzten Jahren zwar deutlich intensiviert, bezieht sich jedoch mehrheitlich auf ältere Kinder. Dazu zählen internationale Vergleichsstudien wie TIMMS (Trends International Mathematics and Science Study) oder PISA (Programme for International Student Assessment). Diese konnten Zusammenhänge von der mathematischen Leistung mit dem sozioökonomischen Status, einem möglichen Migrationshintergrund oder dem Herkunftsland aufzeigen und berichteten teils auch über geschlechterspezifische Leistungsunterschiede. Deutlich weniger Forschungsergebnisse liegen zum Zusammenhang mathematischer Kompetenzen mit Kontextfaktoren von jüngeren Kindern vor. Insbesondere existieren kaum Studien, die den Einfluss von verschiedenen Faktoren auf die mathematische Kompetenzentwicklung im Kindergartenalter beleuchten.

Im Folgenden wird der aktuelle Forschungsstand zum Zusammenhang von mathematischen Kompetenzen mit den Individualfaktoren Vorwissen, Intelligenz, Geschlecht und Alter aufgearbeitet. Danach werden Zusammenhänge mit Merkmalen der sozialen Umwelt wie sozioökonomischer Status, Migrationshintergrund, Erstsprache und elterliches Unterstützungsverhalten aufgezeigt. Fokussiert werden Studien, die Kinder im Vorschulalter untersuchten. Dort, wo keine oder nur sehr wenige vorhanden sind, werden auch Studien aus der Grundschule beigezogen. Das Kapitel endet mit einer Zusammenfassung der wichtigsten Aspekte.

5.1 Vorwissen

Wie in Kapitel 3 dargestellt, leisten frühe mathematische Kompetenzen einen grossen Beitrag zur Erklärung späterer Unterschiede in den mathematischen Kompetenzen. Kinder mit höheren mathematischen Kompetenzen im Kindergarten zeigen tendenziell auch höhere mathematische Leistungen in der Schule. Und Kinder mit niedrigen mathematischen Kompetenzen im Kindergarten haben ein erhöhtes Risiko, in der Schule niedrigere mathematische Leistungen zu erzielen. Das bedeutet, dass mathematische Entwicklungsverläufe von Kindern mit unterschiedlichem mathematischem Vorwissen verschieden verlaufen. Um solche Entwicklungsverläufe zu untersuchen, braucht es aufwendige Längsschnittstudien. Im Bereich Vorschule liegen nur wenige Studien vor, die die Einflussfaktoren auf den mathematischen Entwicklungsverlauf innerhalb der Kindergartenzeit untersuchten. Um dennoch tendenzielle Aussagen über den Einfluss des Vorwissens auf die mathematische Leistungsentwicklung machen zu können, werden auch Studien aus der Grundstufe herangezogen bzw. solche, die den Einfluss des mathematischen Vorwissens im Kindergarten auf schulische mathematische Leistungen in den Blick nehmen. Nachfolgend werden zuerst allgemeine Zusammenhänge zwischen mathematischem Vorwissen und mathematischer Leistungsentwicklung beschrieben. Danach wird auf den Zusammenhang eines niedrigen mathematischen Vorwissens und einer späteren Rechenschwäche eingegangen.

5.1.1 Zusammenhang zwischen mathematischem Vorwissen und mathematischer Leistungsentwicklung

In einer Studie von Ditton und Krüsken (2009) wurden die Schulleistungen in Mathematik von 1249 Schülerinnen und Schülern aus Bayern und Sachsen jeweils zum Schuljahresende der zweiten bis vierten Jahrgangsstufe erhoben. Insgesamt wurden dabei hohe lineare Leistungszuwächse, eine hohe Stabilität und eine reduzierte Leistungsstreuung beim zweiten und dritten Testzeitpunkt offensichtlich. Dabei hatten in beiden Bundesländern Kinder mit den schwächeren Eingangsleistungen über die drei Testzeitpunkte den höchsten Leistungszuwachs und Kinder mit hohen Eingangsleistungen den kleinsten Kompetenzzuwachs. Ähnliche Ergebnisse brachte auch die Studie von Karst und Lipowsky (2013) hervor. Im Rahmen der PERLE-Studie (Leistungs- und Persönlichkeitsentwicklungen von Grundschulkindern) wurden 624 Kinder zu drei Messzeitpunkten (Anfang erster Klasse, Ende erster Klasse und Ende zweiter Klasse) getestet. Insgesamt waren dabei ein signifikanter Effekt des Messzeitpunktes und eine insgesamt lineare Kompetenzentwicklung zu beobachten. Um Unterschiede zwischen Kindern mit verschiedenem Vorwissen zu untersuchen, wurden die Kinder in fünf Leistungsgruppen (Quintile) eingeteilt. Dabei zeigte sich, dass die Leistungen zum ersten Testzeitpunkt noch sehr stark variierten, während sich die Leistungsheterogenität beim zweiten Testzeitpunkt verringerte. Dies ist vor allem auf den Leistungszuwachs des schwächsten Quintils zurückzuführen. Über alle drei Testzeitpunkte gesehen machten auch hier die Kinder mit den niedrigsten Eingangsleistungen die grössten Fortschritte und die leistungsstärksten Schülerinnen und Schüler die kleinsten. Auffällig war, dass die Angleichung des Leistungsniveaus vor allem innerhalb des ersten Schuljahres erfolgte, während im zweiten Schuljahr die Leistungsentwicklungen praktisch parallel verliefen. Auch in der Untersuchung von Fritz et al. (2018) wurde deutlich, dass die Kinder mit den niedrigsten Ausgangsleistungen im Kindergarten acht Monate nach Schulbeginn den grössten Kompetenzzuwachs erreichten, während die Kinder mit der höchsten Ausgangsleistung den geringsten Lernzuwachs aufwiesen. Es wurden die Daten einer Trainingsstudie (Langhorst, Hildenbrand, Ehlert & Fritz, 2013) verwendet, wobei nur die Kinder der Kontrollgruppe in den Analysen berücksichtigt wurden, die keine mathematische Förderung erhielten. Insgesamt konnten die Daten von 62 Kindern verwendet werden. Auch wenn diese drei Studien nachweisen, dass Kinder mit niedrigem mathematischem Vorwissen im Vergleich zu Kindern mit höherem Vorwissen tendenziell einen grösseren Kompetenzzuwachs haben, vermag der Leistungsrückstand nicht aufgeholt werden. Im Gegenteil gibt es Hinweise darauf, dass ein niedriges mathematisches

Vorwissen einen Risikofaktor für eine spätere Rechenschwäche darstellt. Auf diesen Zusammenhang wird im nächsten Abschnitt näher eingegangen.

5.1.2 Zusammenhang zwischen mathematischem Vorwissen und späterer Rechenschwäche

Einige Studien erforschten spezifisch den Zusammenhang von mathematischen Kompetenzen im Kindergarten mit einer späteren Rechenschwäche. In der Studie von Krajewski (2008) mit 148 Kindern wird z. B. darüber berichtet, dass mehr als 60 % der rechenschwachen Kinder in der ersten Klasse und 47 % der rechenschwachen Kinder in der zweiten Klasse bereits im Kindergarten als Risikokinder identifiziert worden sind. Zur Risikoklassifikation wurde jeweils das Mengen- und Zahlenvorwissen herangezogen. Diese Ergebnisse sind vergleichbar mit denen aus den Untersuchungen von von Aster et al. (2007) und von Desoete et al. (2012). Von Aster et al. (2007) untersuchte den Zusammenhang einer Rechenschwäche in der Schule mit den mathematischen Kompetenzen im Kindergarten. So zeigten Kinder der zweiten Klasse mit einer diagnostizierten Rechenstörung bereits im Kindergarten niedrige mathematische Fertigkeiten. Ähnlich verhielt es sich auch bei Desoete et al. (2012). In dieser Studie wurde deutlich, dass Kinder mit sehr schwachen mathematischen Kompetenzen in der zweiten Klasse schon im Kindergarten insgesamt niedrige mathematische Kompetenzen aufwiesen. Es zeigte sich ferner, dass insbesondere die Kombination von niedrigen nicht symbolischen und niedrigen symbolischen Kompetenzen (vgl. Abschnitt 3.3) ein erhöhtes Risiko darstellt, eine Rechenschwäche zu entwickeln. Eine amerikanische Langzeitstudie (Geary, Hoard, Nugent & Bailey, 2013) konnte unter Kontrolle verschiedener demographischer und individueller Faktoren sogar den Zusammenhang von niedrigen mathematischen Kompetenzen im Kindergarten und sehr schwachen mathematischen Leistungen im Alter von 13 Jahren belegen. Die Daten stammten aus einer Längsschnittstudie mit 180 Jugendlichen, die die mathematische Entwicklung im Zusammenhang mit dem Risiko einer Lernbehinderung fokussierte.

Die Ergebnisse legen nahe, dass sich Kinder mit unterschiedlichem mathematischem Vorwissen unterschiedlich entwickeln. Tendenziell verzeichnen Kinder mit niedrigem mathematischem Vorwissen im Kindergarten bzw. der ersten oder zweiten Klasse einen höheren Leistungszuwachs über zwei Jahre als Kinder mit mittlerem oder hohem Vorwissen. Trotzdem haben Kinder mit niedrigen mathematischen Kompetenzen im Kindergarten ein erhöhtes Risiko, eine Rechenschwäche zu entwickeln, als ihre Peers mit höheren mathematischen Kompetenzen.

5.2 Intelligenz

Studien weisen darauf hin, dass das mathematische Vorwissen eine bedeutendere Rolle in der Voraussage von mathematischen Kompetenzen spielt als die Intelligenz. Unter Intelligenz werden hier kognitive Fähigkeiten verstanden, die durch einen normierten Intelligenztest gemessen werden (Stern & Neubauer, 2013). Die grössere Bedeutung des mathematischen Vorwissens im Vergleich mit der Intelligenz konnte auf eindrückliche Art und Weise in der 1980 gestarteten Münchner-LOGIK-Studie nachgewiesen werden. Die neun Jahre vorher erfassten mathematischen Leistungen konnten die Mathematikleistung der elften Klasse besser vorhersagen als die zum gleichen Zeitpunkt erhobene Intelligenz. Der Einfluss der Intelligenz auf die mathematische Kompetenzentwicklung wurde vor allem dadurch ersichtlich, dass sich Kinder mit einer höheren Intelligenz im Verlaufe der Schulzeit mehr mathematisches Wissen aneignen konnten. Eine niedrige Intelligenz konnte dabei durch höhere frühe mathematische Kompetenzen kompensiert werden, niedriges mathematisches Vorwissen durch eine höhere Intelligenz dagegen nicht (Stern, 1998). Trotzdem legen verschiedene Untersuchungen nahe, dass die Intelligenz insgesamt mittelhoch mit den Schulleistungen der Grundschulzeit korreliert (z. B. Bullock & Ziegler, 1997; Weinert & Helmke, 1997) und dass auch die Leistungen in Mathematik mit der Intelligenz zusammenhängen (z. B. Tiedemann & Billmann-Mahecha, 2004). Die Ergebnisse von Studien, die den Einfluss der im Kindergarten gemessenen Intelligenz auf die mathematischen Kompetenzen im Kindergarten oder der ersten bis zweiten Klasse erforschten, sind nicht einheitlich.

Vier Untersuchungen konnten einen Einfluss der Intelligenz auf die mathematischen Kompetenzen im Kindergarten nachweisen (Dornheim, 2008; Hauser et al., 2014; Hornung et al., 2014; Sale, Schell, Koglin & Hillenbrand, 2018). In der Querschnittstudie von Sale et al. (2018) wurden Einflussfaktoren auf mathematische Kompetenzen vor dem Schuleintritt bei 158 Vorschulkinder in Bremen und Niedersachsen untersucht. Die Intelligenz wurde mit dem Grundintelligenztest Skala 1 (CFT 1-R, Weiss & Osterland, 2012) erfasst. Das mathematische Vorwissen wurde nicht in die Untersuchung einbezogen. Bei den anderen drei Studien war der Einfluss der Intelligenz auch unter Einbezug des mathematischen Vorwissens signifikant. Hauser et al. (2014) untersuchten insgesamt 329 sechsjährige Kinder aus 35 Kindergartenklassen zu zwei Testzeitpunkten. Die Intelligenz wurde dabei mit zwei Subtests aus dem sprachfreien Grundintelligenztest CFT 1 (Weiss, Cattell & Osterland, 1997) gemessen. In der Studie von Hornung et al. (2014) erwies sich die Intelligenz im Kindergarten als Prädiktor für die arithmetischen Leistungen in der ersten Klasse. Die Intelligenz wurde hier mit dem nonverbalen Intelligenztest Coloured Progressive Matrices gemessen (CPM; Bullheller & Häcker, 2002). Die dabei erhaltenen Ergebnisse decken sich mit der Studie von Dornheim (2008). Die Intelligenz, gemessen mit drei Subtests aus dem CFT 1 (Weiss et al., 1997), erwies sich als ein signifikanter Prädiktor für die mathematischen Leistungen in der ersten bzw. zweiten Klasse, leistete aber mit 2 bis 10 % Varianzaufklärung im Gegensatz zum Zahlenvorwissen mit einer Varianzaufklärung zwischen 34 und 41 % einen eher geringen Beitrag.

In den Studien von Schuchardt et al. (2014) und Weisshaupt et al. (2006) liess sich kein Einfluss der Intelligenz auf die mathematischen Kompetenzen feststellen. Schuchardt et al. (2014) untersuchten im Rahmen ihrer Längsschnittstudie verschiedene Einflussfaktoren auf die mathematische Entwicklung von 132 Kindern aus 15 Kindergärten im städtischen Kontext. Die Intelligenz wurde hier mit dem nonverbalen Intelligenztest Coloured Progressive Matrices (CPM; Bullheller & Häcker, 2002) erhoben. Die mathematischen Kompetenzen wurden über Aufgaben zu Zählfertigkeiten und Zahleneigenschaften ermittelt. Dabei leistete die Intelligenz keinen bedeutsamen Beitrag zur Varianzaufklärung der mathematischen Kompetenzen ein Jahr später. In der Studie von Weisshaupt et al. (2006) verringerte sich die Rolle der Intelligenz zur Vorhersage von mathematischen Leistungsunterschieden, sobald die Vorkenntnisse der Kinder herangezogen wurden. Die kognitiven Fähigkeiten wurden mit drei Untertests des CFT1 (Weiss et al., 1997) erhoben. Die im Kindergarten ermittelte Intelligenz besass dabei keinen Vorhersagewert für die mathematischen Leistungen am Ende der ersten Klasse.

In zwei Studien war ein indirekter Effekt der Intelligenz beobachtbar (Gallit et al., 2018; Krajewski & Schneider, 2006). In der Studie von Gallit et al. (2018) offenbarten bei der Überprüfung der Vorhersage des Rechnens in der ersten Klasse die erfassten kognitiven Leistungen im Kindergarten indirekte Effekte über das Zahlen- und Mengenvorwissen. Die kognitiven Leistungen wurden mit den Untertests nonverbale und verbale Intelligenz aus dem BUEVA-III von Esser und Wyschkon (2016) gemessen. In der vierjährigen Längsschnittstudie von Krajewski und Schneider (2006) klärte die Intelligenz insgesamt 10 % der Varianz in frühen mathematischen Basisfertigkeiten auf. Diese wurden mit einer Kurzversion des Grundintelligenztests CFT 1 (Weiss et al., 1997) gemessen. Mit dem Einbezug der mathematischen Vorläuferfertigkeiten zur Erklärung der späteren schulischen

Mathematikleistungen war dann aber nur noch ein indirekter Einfluss der Intelligenz, mediiert über Aufgaben zum Kardinalverständnis, ersichtlich (Krajewski & Schneider, 2009).

Insgesamt brachten die zitierten Studien uneinheitliche Ergebnisse zum Einfluss der Intelligenz auf die mathematischen Kompetenzen hervor. Dabei scheinen die eingesetzten Intelligenztests keinen Einfluss auf die Ergebnisse zu haben. Der am häufigsten verwendete Grundintelligenztest CFT 1 (Weiss et al., 1997) kam sowohl in Studien, die einen Einfluss der Intelligenz auf die mathematischen Leistungen nachweisen konnten (Dornheim, 2008; Hauser et al., 2014), als auch in je einer Studie, die keinen (Weisshaupt et al., 2006) oder nur einen indirekten Einfluss (Krajweski & Schneider, 2006) der Intelligenz auf die mathematischen Kompetenzen aufzeigen konnten, zum Einsatz. Auch die mit dem Intelligenztest Coloured Progressive Matrices (CPM; Bullheller & Häcker, 2002) gemessenen kognitiven Fähigkeiten erwiesen sich in einer Studie als signifikanter Prädiktor für die mathematischen Leistungen (Hornung et al., 2014), während dies in der anderen nicht der Fall war (Schuchardt et al., 2014).

Betrachtet man die beiden Studien, die den Einfluss der Intelligenz auf die mathematischen Kompetenzen im Kindergarten, unter Einbezug des mathematischen Vorwissens, in den Blick nahmen, sind zwei unterschiedliche Ergebnisse zu verzeichnen. In der Studie von Hauser et al. (2014) hatte die Intelligenz einen Einfluss, in der Untersuchung von Schuchardt et al. (2014) indes nicht. Allerdings wurden in der Studie von Schuchardt et al. (2014) das visuell-räumliche Arbeitsgedächtnis und die Benenngeschwindigkeit als Prädiktoren eingesetzt und erwiesen sich als signifikant. Es ist anzunehmen, dass diese hoch mit der Intelligenz korrelieren und dass aufgrund von Multikollinearität (Bortz & Schuster, 2016) die Variable Intelligenz nicht signifikant wurde.

5.3 Geschlecht

Die Befunde bezüglich des Zusammenhangs von Geschlecht und mathematischen Leistungen sind insgesamt uneinheitlich. Allerdings zeigen sich bei älteren Kinder zunehmend Vorteile zugunsten der Jungen, wie Ergebnisse aus den grossen internationalen Bildungsstudien zeigen. So ermittelten die PISA-Studien (Programm for Internation Student Assessment) von 2000 bis 2015 im OECD-Durchschnitt durchgängig geschlechtsspezifische Leistungsunterschiede für die 15-Jährigen. Jungen erreichten dabei in allen bisherigen PISA-Erhebungsrunden einen signifikant höheren Wert als Mädchen. Dies gilt auch für die Schweiz und Deutschland (Reiss et al., 2016). Bei TIMSS 2015 (Trends in Internation Mathematics und Science Study) zeigte sich in den untersuchten vierten Klassen, dass in praktisch allen Teilnehmerstaaten, in denen es einen signifikanten Geschlechterunterschied in der Mathematikleistung gab, dieser zugunsten der Jungen ausfiel. Dies war auch in Deutschland der Fall, wobei sich die Differenz als eher geringfügig erwies. Die Schweiz hat an dieser Studie nicht teilgenommen. Allerdings konnte Moser Opitz (2013) in ihrer Studie aus dem Jahr 2007 die höheren Leistungen der Jungen auch für die Schweiz bestätigen. Sie untersuchte 2458 Kinder der fünften und achten Klasse aus der deutschsprachigen Schweiz. Dabei erbrachten die Mädchen signifikant niedrige Mathematikleistungen. Ebenso stellten Winkelmann, Heuvel-Panhuizen und Robitzsch (2008) in einer repräsentativen deutschen Grundschulstudie mit fast 10'000 Schülerinnen und Schülern der dritten und vierten Klasse durchgängig bessere Leistungen der Jungen fest.

Weniger gut erforscht sind indes Geschlechterunterschiede hinsichtlich der mathematischen Leistung von jüngeren Kindern. Die vorliegenden Studien zu dieser Thematik brachten uneinheitliche Ergebnisse für Kindergartenkinder hervor. In den Studien von Rohe und Quaiser-Pohl (2010), Niklas und Schneider (2012), Dornheim (2008), Sahr (2012) sowie Bonny und Lourenco (2013) liess sich mehrheitlich kein Einfluss des Geschlechts auf die mathematischen Leistungen im Kindergarten feststellen. Obwohl es sich bei diesen Untersuchungen um Längsschnittstudien handelt, bezieht sich der Einfluss des Geschlechtes nicht auf den Kompetenzzuwachs, sondern auf die gemessenen mittleren Leistungen zum jeweiligen Testzeitpunkt. Auch in der Querschnittuntersuchung von Sale et al. (2018) hatte das Geschlecht keinen Einfluss auf die mathematischen Kompetenzen von Vorschulkindern. Rohe und Quaiser-Pohl (2010) erforschten unter anderem Geschlechterunterschiede in den mathematischen Vorläuferfertigkeiten bei 146 Kindergartenkindern. Dabei gab es im Kindergarten insgesamt auch keine geschlechterspezifischen Unterschiede. Eine Ausnahme bildete jedoch der Untertest Rechnen. Hier erzielten die Jungen signifikant höher Leistungen, wobei die Effektstärke eher gering war. Niklas und Schneider (2012) konnten in ihrer Studie mit 483 Kindern im Vorschulalter ebenfalls keine Unterschiede in den mathematischen Leistungen zwischen Mädchen und Jungen nachweisen. Allerdings erzielten die Jungen in der ersten Klasse signifikant höhere Werte. Damit verbunden war auch ein höheres mathematisches Selbstkonzept. Dieser Befund bestätigte sich in der Studie von Dornheim (2008), die im Kindergarten ebenfalls noch keine wesentlichen geschlechterspezifischen Unterschiede in den mathematischen Leistungen ausfindig machen konnte. Am Ende der ersten Klasse offenbarte sich dann aber ein höchst signifikanter Unterschied mit hoher Effektstärke zugunsten der Jungen. In einer weiteren Studie untersuchte Sahr (2012) 88 Kinder in der Vorschule sowie der ersten und zweiten Klasse. Im Kindergarten und der zweiten Klasse waren keine geschlechterspezifischen Unterschiede erkennbar. In der ersten Klasse erzielten die Jungen jedoch bessere Leistungen. Auch in der Studie von Bonny und Lourenco (2013) konnten keine mathematischen Leistungsunterschiede zwischen den drei- bis fünfjährigen Mädchen und Jungen nachgewiesen werden.

Andere Studien weisen signifikante Unterschiede bereits im Kindergarten nach (Anders et al., 2012; Krajewski, 2008; Weinhold Zulauf, Schweiter & von Aster, 2003). In der Untersuchung von Weinhold Zulauf et al. (2003) mit insgesamt 334 Kindergartenkindern erzielten die Mädchen beim ersten Testzeitpunkt im ersten Halbjahr noch leicht höhere mathematische Werte. Die Jungen hatten aber einen höheren Leistungszuwachs und schnitten im zweiten Halbjahr signifikant besser ab. Ähnlich verhielt es sich auch bei Anders et al. (2012), die die mathematische Kompetenzentwicklung von 532 Kindern in 97 Vorschulen über drei Messzeitpunkte erforschten. Das durchschnittliche Alter bei der ersten Messung betrug drei Jahre, bei der zweiten vier Jahre und bei der dritten Messung fünf Jahre. Dabei starteten die Mädchen ebenfalls mit höheren mathematischen Kenntnissen, die Jungen zeigten aber einen grösseren Leistungszuwachs. Auch die Studie von Krajewski (2008) konnte querschnittlich nachweisen, dass Jungen im Kindergarten im Gesamtwert Zahlenvorwissen überlegen waren. Die Unterschiede wurden insbesondere im arabischen Zahlwissen und im Zählen ersichtlich. Die Effektstärken waren indes nur mittelhoch. Im Mengenvorwissen wiesen die Jungen tendenziell ebenfalls einen Vorsprung auf, der aber nicht signifikant wurde.

Zusammenfassend weisen die Ergebnisse der Studien darauf hin, dass die Jungen tendenziell mit zunehmendem Alter einen mathematischen Leistungsvorteil gegenüber den Mädchen entwickeln. Bei den 15-Jährigen sind in Deutschland und der Schweiz durchgängig Leistungsunterschiede zu verzeichnen. Zu den Ursachen für diese Geschlechterunterschiede gibt es sowohl Anlage- als auch Umwelttheorien. Da für biologische Unterschiede wenig aussagekräftige Belege vorhanden sind (Ceci, Williams & Barnett, 2009), kann davon ausgegangen werden, dass vor allem soziokulturelle Faktoren hierauf einen Einfluss haben. Relevant erscheinen dahingehend stereotype Einstellungen von Bezugspersonen. Von Jungen werden in der Regel ein grösseres Interesse für und höhere Fähigkeiten in Mathematik erwartet als von Mädchen (Hyde, Lindberg, Linn, Ellis & Williams, 2008). Zudem besitzen Mädchen bei vergleichbarem Leistungsniveau oft auch ein ungünstigeres Fähigkeitsselbstkonzept für Mathematik (Bos, 2008; Schilling, Sparfeldt & Rost, 2006). Diese Unterschiede im Fähigkeitsselbstkonzept zeigen sich nicht erst im höheren Schulalter, sondern bereits im Vorschulalter. Jungen weisen auch auf dieser Stufe tendenziell ein höheres mathematisches Selbstkonzept auf, wobei sich dieser Unterschied im Verlaufe der Grundschulzeit zumeist verfestigt und dann relativ stabil bleibt (Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Marsh, Ellis & Craven, 2002). Dieses höhere mathematische Selbstkonzept der Jungen im Vorschulalter kann jedoch allenfalls zur Erklärung von Unterschieden in den mathematischen Kompetenzen im Kindergarten herangezogen werden. So zeigen doch drei Studien, dass die Jungen bereits im Kindergarten über höhere mathematische Kompetenzen verfügen als die Mädchen. Die Unterschiede im Selbstkonzept könnten wiederum auf die erwähnten Geschlechterstereotypien zurückgeführt werden, die sich im Denken und Handeln von Bezugspersonen ausdrücken (Ehm, Duzy & Hasselhorn, 2011). Dennoch darf nicht vergessen werden, dass die Effekte eher gering waren und dass die Mehrheit der berichteten Studien belegt hat, dass im Kindergarten noch keine geschlechterspezifischen Unterschiede in den mathematischen Kompetenzen vorhanden sind.

5.4 Alter

Wie in Kapitel 4 beschrieben, beginnt das mathematische Lernen nicht erst in der Schule oder im Kindergarten. Mathematische Kompetenzen entwickeln sich ab der frühsten Kindheit. Kleinste Kinder nehmen bereits Veränderungen in Mengen wahr. Mit dem Beginn des Sprechens lernen sie einzelne Zahlwörter kennen und erwerben mit der Zeit die Zahlwortreihe. Damit verbunden gelingt allmählich auch das exakte Bestimmen einer Anzahl. Kinder haben also bereits im Alter von fünf Jahren eine beachtliche Anzahl von mathematischen Kompetenzen erworben. Mit zunehmendem Alter vergrössern und vertiefen sich diese mathematischen Kompetenzen. Es kann deshalb davon ausgegangen werden, dass das Alter insgesamt ein bedeutender Prädiktor für die Vorhersage von mathematischen Kompetenzen ist. Je älter Kinder sind, desto höher sind in der Regel ihre mathematischen Kompetenzen. Betrachtet man aber eine kürzere Altersspanne, in diesem Fall den Einfluss des Alters auf die mathematischen Kompetenzen von Kindergartenkindern, sind die Ergebnisse nicht einheitlich. Die Studien von Jörns, Schuchardt, Mähler und Grube (2013), Anders et al. (2012) und Sale et al. (2018) beleuchten den Einfluss des Alters im Vorschulalter. Bei Jörns et al. (2013) hatten die älteren Kinder höhere mathematische Kompetenzen als die jüngeren Kinder. Sie untersuchten insgesamt 142 Kinder im Alter von vier bis fünf Jahren. Auch bei Sale et al. zeigten die älteren Kinder höhere Kompetenzen. Die Studie von Anders et al. (2012) erforschte mittels Wachstumskurvenmodell unter anderem auch den Einfluss des Alters auf die Kompetenzentwicklung von Vorschulkindern. Dabei zeigte sich, dass die älteren Kinder sowohl höhere mathematische Leistungen beim ersten Testzeitpunkt als auch einen grösseren mathematischen Kompetenzzuwachs aufwiesen. In anderen Studien wurde der Einfluss des Alters im Kindergarten auf die mathematische Kompetenzentwicklung in der Grundschule in den Fokus genommen. In der Studie von Manfra et al. (2017) erwies sich das Alter der Kinder im Kindergarten als ein signifikanter Prädiktor für die Vorhersage der mathematischen Leistungen in der dritten Klasse. Es waren die jüngeren Kinder, die in der dritten Klasse insgesamt die höheren Leistungen zeigten. In der Studie von Toll et al. (2016) war das Alter der Kinder im Kindergarten ebenfalls ein Prädiktor für die mathematischen Leistungen in der ersten Klasse. Auch hier besassen die jüngeren Kindergartenkinder in der Schule höhere Kompetenzen; allerdings nur im Bereich Faktenabruf von Additions- und Subtraktionsaufgaben und nicht bei Aufgaben zum mathematischen Problemlösen. In der erwähnten Studie von Rohe und Quaiser-Pohl (2010) erwies sich das Alter im Kindergarten nur bei den Mädchen als statistisch bedeutsamer Prädiktor zur Vorhersage der Mathematikleistung in der zweiten Klasse. Je jünger die Mädchen waren, desto höher waren die Leistungen.

Die Ergebnisse zeigen, dass ältere Kinder mit höheren mathematischen Kompetenzen in den Kindergarten eintreten, dass aber die jüngeren Kinder im Kindergarten tendenziell in der Schule höhere Leistungen aufweisen. Es ist denkbar, dass die jüngeren Kinder durch die pädagogischen Fachkräfte intensiver gefördert wurden, um ihren anfänglichen Rückstand aufzuholen, und sich diese Förderung positiv auf die Leistungsentwicklung ausgewirkt hat. Möglich ist auch, dass die jüngeren Kinder den älteren nacheifern und sich intensiver mit mathematischen Themen auseinandersetzen. Die Ergebnisse und Interpretationen sind allerdings mit Vorsicht zu betrachten. Denn mit Blick auf die Leistungsentwicklung innerhalb des Kindergartens wird innerhalb der Studie von Anders et al. (2012) deutlich, dass die älteren Kinder hier auch einen grösseren Kompetenzzuwachs hatten.

5.5 Merkmale der sozialen Umwelt

Neben den oben beschriebenen individuellen Faktoren legen verschiedene Forschungsarbeiten auch nahe, dass eine mögliche Ursache für die grossen mathematischen Unterschiede in den

Merkmalen der sozialen Umwelt der Kinder zu finden ist (z. B. Schuchardt et al., 2014). Dazu zählen soziökonomischer Status, eventueller Migrationshintergrund, Erstsprache und elterliches Unterstützungsverhalten. Im Folgenden wird der Forschungsstand zum Zusammenhang dieser kontextuellen Faktoren mit mathematischen Kompetenzen aufgezeigt.

5.5.1 Sozioökonomischer Status

In Bezug auf ältere Kinder war im Rahmen der TIMSS- und der PISA-Studien ein durchweg signifikanter Leistungsvorsprung in Mathematik bei Kindern aus Familien mit einem höheren soziökonomischen Status beobachtbar (Reiss et al., 2016; Wendt et al., 2016). Dieselben Ergebnisse lieferten auch Studien zum Einfluss des sozioökonomischen Status auf die mathematischen Kompetenzen im Kindergarten. In der Studie von Schuchardt et al. (2014) wurde der sozioökonomische Status als Summenwert über den Schul- und Berufsabschluss der Eltern erhoben. Dabei zeigte sich ein signifikanter Einfluss der Bildung der Eltern auf die mathematischen Fähigkeiten der 132 fünfjährigen Kinder. Bei Lehrl, Kuger und Anders (2014) hatte der sozioökonomische Status bei Kindern ohne Migrationshintergrund einen Einfluss auf die mathematische Kompetenzentwicklung im Kindergarten. Kinder aus Familien mit höherem sozioökonomischem Status erzielten einen grösseren Kompetenzzuwachs. Auch bei Anders et al. (2012) beeinflusste der sozioökonomische Status sowohl die Ausgangsleistung als auch den Kompetenzzuwachs der mathematischen Leistungen bei Kindern im Alter zwischen drei und fünf Jahren signifikant.

5.5.2 Migrationshintergrund

Sowohl bei der TIMSS- als auch bei der PISA-Studie liess sich ein signifikanter Einfluss des Migrationshintergrunds auf die mathematischen Kompetenzen von 10- bis 15-jährigen Jugendlichen feststellen. Schülerinnen und Schüler mit Migrationshintergrund zeigten dabei geringere Leistungen als ihre Peers ohne Migrationshintergrund (Reiss et al., 2016; Wendt et al., 2016). Bei Kindern im Vorschulalter beeinflusste der Migrationshintergrund die mathematischen Kompetenzen (Jörns et al., 2013; Schuchardt et al., 2014) bzw. den Kompetenzzuwachs (Lehrl et al., 2014) jedoch offenbar nicht.

Gleichwohl sind diese Ergebnisse mit Vorsicht zu betrachten. Erstens ist die Operationalisierung der Variable Migrationshintergrund nicht in allen Studien gleich. Bei TIMSS und PISA wurde die Variable Migrationshintergrund vergeben, wenn das Geburtsland der Eltern nicht Deutschland (oder die Schweiz) war. Bei Schuchardt et al. (2014) wurde der Status Migrationshintergrund erst vergeben, wenn der Geburtsort der Eltern im Ausland und die Familiensprache nicht Deutsch war. In der Studie von Jörns et al. (2013) bedeutete die Variable Migrationshintergrund, wenn mindestens ein Elternteil nicht in Deutschland geboren wurde und/oder die Muttersprache nicht Deutsch war. Bei Lehrl et al. (2014) musste mindestens ein Elternteil eine andere Erstsprache als Deutsch sprechen, damit der Status Migrationshintergrund zutraf. Zweitens ist der Einfluss des Migrationshintergrunds auf die mathematischen Kompetenzen auch von anderen Variablen abhängig, die in die Analysen mitaufgenommen wurden. Prädiktoren wie Erstsprache, sozioökonomischer Status oder Sprachverständnis korrelieren oft stark mit dem Kriterium Migrationshintergrund. Aufgrund von Multikollinearität können so einzelne Variablen nicht signifikant sein, obwohl diese in der Realität einen Einfluss haben (Bortz & Schuster, 2016). Wurden keine mit dem Migrationshintergrund korrelierende unabhängige Variablen ins Modell aufgenommen, zeigt sich z. B. in der Studie von Heinze, Herwartz-Emden und Reiss (2007), dass Kinder mit Migrationshintergrund am Ende des ersten Schuljahres signifikant niedrige mathematische Leistungen hatten. Wurden allerdings die Intelligenz und der Sprachstand kontrolliert, liessen sich keine Unterschiede mehr zwischen Kindern mit und ohne Migrationshintergrund beobachten. Es kann also davon ausgegangen werden, dass auch bei jüngeren Kindern der Migrationshintergrund einen Einfluss auf die mathematischen Kompetenzen haben kann, dass aber korrelierende Variablen wie Erstsprache oder sozioökonomischer Status bedeutender sind.

5.5.3 Erstsprache

Verschiedene Studien haben Zusammenhänge zwischen sprachlichen und mathematischen Kompetenzen aufgezeigt (Prediger, Wilhelm, Büchter, Gürsoy & Benholz, 2015; Rösch & Paetsch, 2011; Ufer & Mehringer, 2013). Mit Blick auf jüngere Kinder ist die Studie von Heinze et al. (2007) von Bedeutung. Darin konnten die Autoren nachweisen, dass die Sprache für den Erwerb mathematischer Kenntnisse im ersten Schuljahr von entscheidender Bedeutung ist. Insbesondere gab es Hinweise darauf, dass die Sprache den Aufbau mentaler Repräsentationen beeinflusst. Weniger betroffen von der Sprache waren hingegen Aufgaben zu den Grundoperationen in der Gleichungsschreibweise. Hinsichtlich des Zusammenhangs zwischen sprachlichen und mathematischen Kompetenzen im vorschulischen Bereich konnte die Studie von Jörns et al. (2013) belegen, dass Kindergartenkinder mit niedrigem Sprachverständnis auch niedrige mathematische Kompetenzen aufwiesen. In der Untersuchung von Sale et al. (2018)

zeigte sich ebenfalls ein signifikanter positiver Zusammenhang zwischen mathematischen und sprachlichen Kompetenzen im Kindergarten.

Zum spezifischen Einfluss der Erstsprache, und nicht generell von sprachliche Kompetenzen, auf die mathematischen Leistungen liegen insgesamt nur wenige Forschungsarbeiten vor. Dies ist jedoch auch dem Umstand geschuldet, dass die Variable oft im Zusammenhang mit dem Migrationshintergrund vergeben wird (vgl. Abschnitt 5.5.2). Bei der Untersuchung von Moser Opitz (2013) erzielten Kinder der fünften Klasse mit Deutsch als Erstsprache bessere Mathematikleistungen als Kinder mit einer anderen Erstsprache. Im achten Schuljahr hingegen konnte kein solcher Unterschied mehr gefunden werden. Es ist zu vermuten, dass Kinder im achten Schuljahr mit einer anderen Erstsprache als Deutsch schon länger in der Schweiz leben als die Mehrheit der Kinder in der fünften Klasse und so mehr Zeit für den Spracherwerb blieb. Bezüglich jüngerer Kinder fand eine Untersuchung mit 355 Kindern in der deutschsprachigen Schweiz heraus, dass Kinder mit einer anderen Erstsprache als Deutsch geringere verbale Zählkompetenzen aufwiesen als deutschsprachige Kinder (Moser Opitz et al., 2010). Da die Zählkompetenzen einen wesentlichen Prädiktor für mathematische Leistungen darstellen (Jordan et al., 2007), dürften Kinder mit einer anderen Erstsprache als Deutsch benachteiligt sein. In der Studie von Anders et al. (2012) zeigten sich Unterschiede im Ausgangsniveau und im Leistungszuwachs der mathematischen Leistung der Kinder, wenn beide Elternteile Deutsch nicht als Muttersprache hatten. Wenn nur ein Elternteil eine andere Muttersprache besass, war der Einfluss nur im Ausgangsniveau ersichtlich, nicht aber im Leistungszuwachs. Bei Sale et al. (2018) hatte die Familiensprache einen Einfluss auf die mathematischen Kompetenzen der zweiten Ebene des ZGV-Modells nach Krajewski (2008), nicht aber auf die erste Ebene (vgl. Abschnitt 4.2.1). Das Autorenteam vermutete, dass die mathematischen Kompetenzen auf Ebene 2 „stärker vom mathematischen Wortschatz und damit indirekt von der gesprochen Sprache zu Hause abhängen" (Sale et al. 2018, S. 381). Insgesamt wird also deutlich, dass ein Einfluss der Erstsprache auf die mathematischen Leistungen vorhanden ist und dass Kinder mit Deutsch als Erstsprache einen Vorteil haben.

5.5.4 Elterliches Unterstützungsverhalten

In engem Zusammenhang mit dem sozioökonomischen Status, dem Migrationshintergrund und der Erstsprache wird auch die Bedeutung des elterlichen Unterstützungsverhaltens für die mathematische Entwicklung diskutiert. Darunter werden eine Vielzahl an Verhaltensweisen, Einstellungen und Haltungen wie gemeinsames Spielen und Lesen, gemeinsame Unternehmungen oder an die Kinder herangetragene Erwartungshaltungen verstanden. Rindermann und Baumeister (2015) konnten in einer Metaanalyse zeigen, dass das Erziehungsverhalten der Eltern für die Erklärung von Unterschieden in der kognitiven Entwicklung von zwei-, drei- und neunjährigen Kindern bedeutsamer ist als der sozioökonomische Status. Der Einfluss des elterlichen Erziehungsverhaltens (home learning environment, HLE) auf die Lese- und Rechenfähigkeiten in der Vorschule (mit drei Jahren), beim Schuleintritt (mit fünf Jahren) und am Ende des dritten Schuljahres (mit sieben Jahren) wurde in einer grossen englischen Studie mit über 2300 Kindern von Melhuish et al. (2008) erforscht. Das wirksame Erziehungsverhalten im Vorschulalter bestand dabei aus elterlichem Vorlesen, Bibliotheksbesuchen, vielfältigem Spiel mit Zahlen und Wörtern, Malen und Zeichen und der Förderung des Erwerbs von Buchstaben, Zahlen, Liedern und Reimen. Kinder aus Familien, in denen häufiger solche Erziehungsaktivitäten angewendet wurden, hatten signifikant höhere mathematische und sprachliche Kompetenzen. Dass der Einfluss der elterlichen Unterstützung für sprachliche und mathematische Kompetenzen von Bedeutung ist, wiesen auch Anders et al. (2012) in ihrer Studie nach. Die Forschungsgruppe untersuchte unter anderem den Einfluss der häuslichen Lernumgebung auf Kompetenzen im Bereich Sprache und Mathematik. Deutlich wurde, dass die häusliche Lernumgebung sowohl im sprachlichen als auch im mathematischen Bereich einen Einfluss auf die Ausgangsleistung der Kinder beim ersten Testzeitpunkt hatte. Kinder, deren häusliche Lernumgebung anregender war, zeigten im Alter von drei Jahren höhere sprachliche und mathematische Kompetenzen. Hingegen hatte die häusliche Lernumgebung keinen Einfluss auf die Kompetenzentwicklung, weder im Bereich Sprache noch im Bereich Mathematik. Andere Ergebnisse fanden sich indes in der Studie von Schuchardt et al. (2014). Hier hatte das Home Numeracy Environment (HLE) einen Einfluss auf die Kompetenzentwicklung der Kindergartenkinder. Je anregender und spezifischer die häusliche Lernumgebung im Bereich Mathematik ausgestaltet war, desto höher waren die Kompetenzen der Kinder zum zweiten Testzeitpunkt. Die häusliche Lernumgebung im Bereich Sprache hatte hingegen keinen Einfluss. Die Forschergruppe kam zu dem Schluss, dass die mathematische Entwicklung nicht durch allgemein häusliche Anregungen gefördert wird, sondern vielmehr durch spezifische Lernerfahrungen mit zahlen- und mengenbezogenen Aktivitäten. Diese Ergebnisse wurden in der Untersuchung von Ceulemans et al. (2017) bestätigt. Auch hier gab es einen positiven Zusammenhang zwischen mathematischen

Anregungen zu Hause und der mathematischen Leistung im Kindergarten. Allerdings war die Stichprobe mit 31 untersuchten Kindern relativ klein.

5.6 Zusammenfassung

In diesem Kapitel wurde aufgezeigt, wie die mathematischen Kompetenzen von Kindern mit individuellen und kontextuellen Faktoren zusammenhängen. Dabei spielt das Vorwissen eine zentrale Rolle. Kinder mit höheren mathematischen Kompetenzen im Kindergarten erzielen tendenziell höhere mathematische Kompetenzen in der Schule. Gleichzeitig haben Kinder mit niedrigem mathematischem Vorwissen ein erhöhtes Risiko, in der Schule eine Rechenschwäche zu entwickeln als ihre Peers mit höheren vorschulischen mathematischen Kompetenzen. Bezüglich der Intelligenz legen verschiedene Studien nahe, dass die kognitiven Fähigkeiten in einem Zusammenhang mit den mathematischen Leistungen von Kindern und Jugendlichen stehen. Im Bereich Kindergarten sind die Forschungsergebnisse zum Einfluss der Intelligenz uneinheitlich. Generell zeigt sich, dass unter Einbezug des mathematischen Vorwissens der Einfluss der Intelligenz sinkt. Zum Einfluss des Geschlechts auf die mathematischen Leistungen ist, über alle Schulstufen betrachtet, insgesamt ein Vorteil zugunsten der Jungen zu verzeichnen, wobei die Ergebnisse im Kindergarten auch hier nicht einheitlich sind. Tendenziell gibt es im Kindergarten noch keine oder nur sehr geringe Geschlechterunterschiede. Bezüglich des Einflusses des Alters liegen Hinweise darauf vor, dass die jüngeren Kinder im Kindergarten später höhere mathematische Kompetenzen zeigen, aber auch, dass ältere Kinder im Kindergarten einen grösseren Kompetenzzuwachs haben. Neben individuellen Faktoren des Kindes liefern auch die primären Sozialisationserfahrungen im Elternhaus Erklärungen für die unterschiedlich ausgeprägten Kompetenzen. Verschiedene Studien konnten nachweisen, dass Kinder aus Familien mit niedrigem sozioökonomischem Status, mit Migrationshintergrund und Kinder mit einer anderen Erstsprache als Deutsch bereits im Kindergarten niedrige mathematische Kompetenzen aufweisen als ihre Peers ohne soziale Benachteiligungen. Zudem gibt es deutliche Hinweise darauf, dass das elterliche Unterstützungsverhalten die mathematischen Kompetenzen der Kinder beeinflusst. Kinder aus Familien, die mehr zahlen- und mengenbezogene Aktivitäten anbieten, haben tendenziell höhere mathematische Kompetenzen. Um die Chancengerechtigkeit von sogenannten benachteiligten Kindern aus Familien mit niedrigem sozioökonomischem Status, mit Migrationshintergrund und mit wenig häuslicher Lernanregung zu erhöhen, wird die Aufgabe des Ausgleichs solcher sozialen Disparitäten in vorschulischen Einrichtungen diskutiert (Becker, 2010).

Nachdem in den Kapiteln 3 bis 5 vorschulische mathematische Kompetenzen, deren Bedeutung für mathematische Leistungen in der Schule, deren Entwicklung und deren Zusammenhänge mit individuellen und kontextuellen Faktoren im Zentrum der Auseinandersetzung standen, widmet sich das nächste Kapitel nun der Messung mathematischer Kompetenzen.

6 Messung mathematischer Kompetenzen im Längsschnitt

Kompetenz wird in dieser Arbeit als das Ergebnis eines leistungsbezogenen Tests verstanden (vgl. Abschnitt 3.1). Dieses Ergebnis wird von verschiedenen individuellen Merkmalen wie Motivation und Wille (Weinert, 2001), aber auch von externen Variablen wie Testarrangement, Testdurchführung oder Testleitung beeinflusst. In der Testtheorie wird in diesem Zusammenhang von Messfehlern gesprochen (Rost, 2004). Die genaue Messung von Kompetenzen stellt deshalb generell eine methodische Herausforderung dar. Wenn zusätzlich die Messung von Veränderungen über die Zeit erfasst werden soll, erhöhen sich die Messfehler durch die mehrfache Messung (Ittel & Merkens, 2006). Zudem setzt die Darstellung von Veränderungen voraus, dass zwischen den Testzeitpunkten Differenzen feststellbar werden. Diese Differenzwerte sind jedoch ebenfalls nicht unproblematisch, da auch sie mit Messfehlern behaftet sind (Rost, 2004). Bei einer längsschnittlichen Messung muss darüber hinaus sichergestellt werden, dass die Ankeritems zu allen Messzeitpunkten und in allen Probandengruppen das gleiche Konstrukt messen. Es ist deshalb erforderlich, verlässliche Erhebungsinstrumente einzusetzen, die testtheoretische Kriterien erfüllen. In diesem Kapitel werden zuerst Axiome und Gütekriterien der klassischen Testtheorie besprochen. Danach werden die Item-Response-Theorie und das darauf basierende Rasch-Modell mit Annahmen, Parameterschätzung und Modellüberprüfungstest erläutert. Des Weiteren werden vier Erhebungsinstrumente zur mathematischen Messung im Kindergartenalter vorgestellt und im Hinblick auf die Möglichkeit zur längsschnittlichen Erfassung diskutiert. Eine Zusammenfassung schliesst das Kapitel ab.

6.1 Klassische Testtheorie

Die Klassische Testtheorie (KTT) stellt den theoretischen Hintergrund zur Konstruktion und Interpretation vieler psychodiagnostischer Tests dar (Moosbrugger & Kelava, 2012). Die Prinzipien der KTT gehen auf Gulliksen (1950) und Novick (1966) zurück und basieren auf Konstrukten, die seit Beginn des 20. Jahrhunderts in der Psychologie zur Messung von Merkmalsunterschieden zwischen Personen beigezogen werden. Die KTT geht davon aus, dass sich der Messwert aus der wahren, tatsächlichen Ausprägung des mit dem Item erfassten Merkmals und einem zufälligen Messfehler zusammensetzt. In fünf Axiomen werden testtheoretische Grundannahmen der KTT formuliert (Döring & Bortz, 2016), in denen wesentliche Annahmen über das wahre Merkmal und den Messfehler getroffen werden. Damit lässt sich die Genauigkeit der Messung einschätzen. Folgende Definitionen beziehen sich auf die Ausführungen in Döring und Bortz (2016), S. 463f.

1. Existenzaxiom: Der wahre Wert existiert als Erwartungswert der Messungen eines Probanden.
2. Verknüpfungsaxiom: Jeder Testwert setzt sich aus einem wahren Merkmalsanteil und einem zufälligen Messfehleranteil zusammen.
3. Unabhängigkeitsaxiom: Es gibt keine Korrelation zwischen den Messfehlern und den wahren Werten bei beliebigen Personen und beliebigen Items.
4. Unabhängigkeit des Messfehlers zwischen Items: Die Fehlerwerte zweier Messungen mit beliebigen Items für dieselbe Person sind unkorreliert. Die Items müssen so konstruiert sein, dass eine Aufgabe unabhängig vom Erfolg oder Misserfolg anderer Testaufgaben gelöst werden kann.
5. Unabhängigkeit des Messfehlers zwischen Personen: Die Fehlerwerte zweier Messungen mit beliebigen Personen mit demselben Item sind unkorreliert.

Auf Basis dieser Axiome können drei zentrale Gütekriterien formuliert werden, die die Qualität eines Tests angeben: Objektivität, Reliabilität und Validität. Ein Test ist dann objektiv, wenn das gemessene Merkmal unabhängig vom Testleitenden oder Testauswertenden ist. Dem Testdurchführenden darf kein Verhaltensspielraum bei der Durchführung, Auswertung und Interpretation eingeräumt werden (Moosbrugger & Kelava, 2012). Unter Reliabilität wird die Zuverlässigkeit des Tests verstanden. Dieser ist dann reliabel, wenn er das zu interessierende Merkmal möglichst genau und ohne Messfehler misst. Unterschieden wird dabei zwischen Retest-, Paralleltest- und Testhalbierungs-Reliabilität sowie interner Konsistenz (Döring & Bortz, 2016). Zur Bestimmung der Retest-Reliabilität wird der Test derselben Stichprobe zweimal vorgelegt. Je höher die Ergebnisse korrelieren, desto höher ist die Retest-Reliabilität. Paralleltest-Reliabilität liegt dann vor, wenn zwei parallele Testformen trotz nicht identischer (sich aber inhaltlich möglichst ähnelnder) Items zu gleichen Testwerten führen. Wenn der Test nicht wiederholt werden kann und auch keine parallele Testform vorliegt, kann die Testhalbierungs-Reliabilität überprüft werden. Dazu wird der Test in zwei Testhälften geteilt und deren Korrelation bestimmt. Die Bestimmung der Reliabilität nach der Testhalbierungsmethode hängt von der Art der zufälligen Halbierung ab. Stabilere Schätzungen ergeben sich durch die Berechnung der internen Konsistenz. Diese besagt, wie stark die einzelnen Items eines Tests untereinander korrelieren. Das Mass für diese Reliabilität in der KTT ist üblicherweise Cronbachs Alpha. Diese Masszahl geht auf Lee Cronbach zurück (Cronbach, 1951) und kann Werte von minus unendlich bis 1 annehmen, wobei nur positive Werte sinnvoll interpretiert werden können. In der Literatur lassen sich verschiedene Angaben darüber finden, welcher Alpha-Wert als ausreichend angesehen wird. Döring und Bortz (2016) nennen α=.80 als anzustrebenden Wert, während Schmitt (1996) α=.70 als üblichen Schwellenwert bezeichnen. Allgemein werden Werte über α=.90 als hoch angesehen (Weise, 1975). Neben der Objektivität und der Reliabilität wird die Validität als drittes Gütekriterium von Tests genannt. Die Validität zeigt an, ob ein Test das misst, was er messen soll. Im Vergleich zu Objektivität und Reliabilität gestaltet sich die Überprüfung der Validität aufwendiger. Es wird zwischen Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität unterschieden. Inhaltsvalidität ist gegeben, wenn die Testitems das zu messende Konstrukt in den zentralen Aspekten mehrheitlich erschöpfend zu messen vermögen. So sollte ein Test zur Erfassung mathematischer Kompetenzen im Kindergarten möglichst alle in den Entwicklungsmodellen zur Zahlbegriffsentwicklung postulierten Kompetenzen überprüfen. Die Höhe der Inhaltsvalidität kann nicht numerisch bestimmt werden, sondern beruht auf subjektiven Einschätzungen (Schnell, Hill & Esser, 1999). Die Kriteriumsvalidität ist dann erfüllt, wenn ein Test ein interessierendes Merkmal so misst, dass es mit einem für das Merkmal relevanten Aussenkriterium übereinstimmt (Döring & Bortz, 2016). Werden beispielsweise die mathematischen Kompetenzen erfasst, so sollten die Testergebnisse mit den mathematischen Leistungen in der Schule (Aussenkriterium) übereinstimmen (beispielsweise mit der Note im Fach Mathematik). Konstruktvalidität ist dann erfüllt, wenn der Testwert inhaltlich und theoretisch begründet hypothesenkonform mit anderen theoretischen Konstrukten korreliert.

Neben der klassischen Testtheorie existiert heute eine neuere Testtheorie, mit der die KTT vorteilhaft ergänzt werden soll: die Item-Response-Theorie (Moosbrugger & Kelava, 2012).

6.2 Item-Response-Theorie

Die Klassische Testtheorie (KTT) geht von der Annahme aus, dass das Ergebnis eines Tests eine mit Messfehlern behaftete Merkmalsausprägung repräsentiert. Die probabilistische Testtheorie oder Item-Response-Theorie (IRT) hingegen sieht das Testergebnis als Indikator latenter Dimensionen. Die IRT formuliert Modelle zu den Zusammenhängen zwischen latenten

Persönlichkeitseigenschaften und den Antworten des Tests. Es interessieren also Wahrscheinlichkeiten für die Lösung von Items in Abhängigkeit von den Fähigkeiten der untersuchten Person und den Schwierigkeiten der Items (Döring & Bortz, 2016). Unterscheiden sich zwei Personen hinsichtlich ihrer Fähigkeiten, wird ein Item von der Person mit höheren Fähigkeiten mit grösserer Wahrscheinlichkeit gelöst werden als von der Person mit niedrigen Fähigkeiten. Ausserdem wird eine Person mit einer bestimmten Fähigkeit von zwei Items das leichtere mit einer höheren Wahrscheinlichkeit lösen als das schwerere. Eine Stärke von IRT-Modellen liegt darin, dass die Aufgabenschwierigkeiten und Personenfähigkeiten auf einer gemeinsamen Skala abgebildet werden (Hartig & Frey, 2013). „Dies ist eine wesentliche Grundlage für die kriteriumsorientierte Definition und Beschreibung von Kompetenzniveaus. Die gemeinsame Skala ist in der diagnostischen Praxis Voraussetzung für niveaubezogene Rückmeldungen von Testergebnissen [...]“ (Hartig & Frey, 2013, S. 2). Die IRT umfasst zahlreiche messtheoretische und psychologische Modelle. Ein einfaches und weit verbreitetes IRT-Modell ist das Rasch-Modell für dichotome Variablen (Hartig & Kühnbach, 2006).

6.2.1 Das Rasch-Modell

Das Rasch-Modell geht auf den dänischen Statistiker Georg Rasch (Rasch, 1960) zurück. Die Grundannahme des Modells besteht darin, dass die Wahrscheinlichkeit, ein Item zu lösen, ausschliesslich von der Fähigkeit der Person und der Schwierigkeit des Items abhängt (Bond & Fox, 2015). Die Art der Beziehung, die die Lösungswahrscheinlichkeit eines Items mit den Fähigkeiten der Person verknüpft, wird Itemcharakteristik (Item Characteristic Curve: ICC) genannt. Der Kurvenverlauf entspricht einer logistischen Funktion. Jedes Item im Test besitzt eine eigene ICC-Kurve.

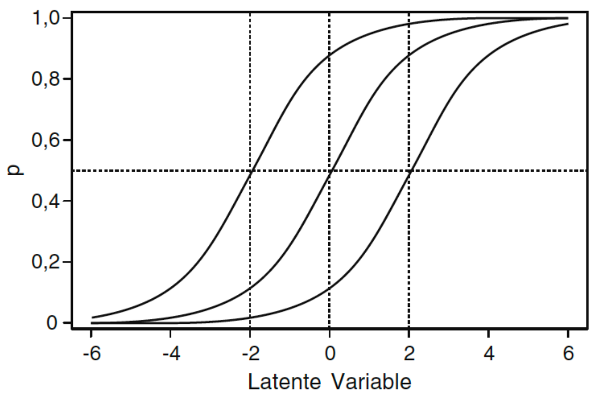

Abbildung 3: Itemcharakteristiken des Rasch-Modells nach Schnell et al. 1999, S. 191

Auf der x-Achse ist die Personenfähigkeit und auf der y-Achse die Lösungswahrscheinlichkeit abgebildet. Diese liegt zwischen 0 und 1. Mit zunehmender Fähigkeit steigt die Wahrscheinlichkeit, dass eine Person das Item lösen kann. Der Wendepunkt liegt jeweils bei einer Wahrscheinlichkeitslösung von 50 %. Hieran lässt sich ablesen, welche Fähigkeit eine Person haben muss, um eine Aufgabe mit einer 50-prozentigen Wahrscheinlichkeit zu lösen. Wenn das Rasch-Modell für einen Test gilt, müssen alle Item-Kurven parallel verlaufen, also dieselbe Steigung haben. Dies ist in der Modellgleichung angelegt. Es gibt keinen Parameter, der die Steigung der Kurve ausdrückt. Die Kurven der einzelnen Items variieren nur in der Verschiebung nach rechts (schwierigeres Item) bzw. nach links (einfacheres Item). Die mittlere Steigung zeigt die Trennschärfe an, die entsprechend für alle Items gleich ist (Rost, 2004). Diese konstante Trennschärfe stellt eine zentrale Eigenschaft des Rasch-Modells dar. Daneben existieren noch weitere Eigenschaften, nämlich die Annahme der Eindimensionalität, der lokalen stochastischen Unabhängigkeit, der spezifischen Objektivität und der Suffizienz der Summenscores (Bond & Fox, 2015). Eindimensionalität meint, dass die Beantwortung eines Items hauptsächlich durch die Fähigkeitsausprägung im Merkmal (z. B. numerische Kompetenz) zustande kommt. Andere Fähigkeiten, wie beispielsweise sprachliche Kompetenz, spielen für die Beantwortung der Items indes keine bedeutende Rolle. In diesem Zusammenhang wird auch von Itemhomogenität gesprochen (Koller, Alexandrowicz & Hatzinger, 2012). Lokale stochastische Unabhängigkeit zwischen Items bedeutet, dass die Wahrscheinlichkeit, ein Item zu lösen, nur vom jeweiligen Item selbst abhängt und nicht von der Lösungswahrscheinlichkeit eines anderen (Strobl, 2015). Diese Eigenschaft kann beispielsweise verletzt sein, wenn Items aufeinander aufbauen und das Lösen des ersten Items die Lösungswahrscheinlichkeit des zweiten Items erhöht. Die Eigenschaft der lokalen stochastischen Unabhängigkeit ist auch dann verletzt, wenn in einem Test zwei inhaltlich sehr ähnliche Items vorkommen. Wenn eine Person das eine Item lösen kann, wird sie mit hoher Wahrscheinlichkeit auch das andere lösen können (Koller et al., 2012). Unter spezifischer Objektivität wird verstanden, dass Aussagen über die Fähigkeit von zwei Personen nicht davon abhängen, anhand welcher Aufgabe sie verglichen werden, und dass umgekehrt die Itemschwierigkeit nicht von der Personenfähigkeit abhängig ist. In diesem Kontext wird auch der Begriff Stichprobenunabhängigkeit oder Subgruppeninvarianz verwendet. Aufgaben müssen für unterschiedliche Personengruppen mit gleicher Fähigkeit gleich schwer zu beantworten sein (Strobl, 2015). Wenn alle beschriebenen Eigenschaften erfüllt sind, dann stellen Personenscores für die Personenfähigkeit und die Itemscores für die Itemschwierigkeiten suffiziente Statistiken dar. Personen mit demselben Personenscore haben dieselbe Fähigkeit, egal welche Items sie gelöst haben. Daneben gilt genauso, dass Items mit demselben Itemscore, unabhängig davon, welche Personen sie gelöst haben, dieselbe Itemschwierigkeit haben (Koller et al., 2012).

Wenn das Rasch-Modell gilt, müssen diese Annahmen für die erhobenen Daten zutreffen. Dabei stehen verschiedene Tests zur Überprüfung zur Verfügung (vgl. Abschnitt 6.2.3). Im Folgenden wird nun aber zuerst auf die Parameterschätzung im Rasch-Modell eingegangen.

6.2.2 Schätzung der Modellparameter

Im Rasch-Modell hängt die Wahrscheinlichkeit, ein Item zu lösen, von der Itemschwierigkeit und der Personenfähigkeit ab (Bond & Fox, 2015). Diese beiden Parameter werden im Rasch-Modell mittels dem in der Statistik weit verbreiteten Konzept der Likelihood geschätzt. Auf die verschiedenen Verfahren der Schätzung (vgl. Rost, 2004, S. 309–317) wird hier jedoch nicht weiter eingegangen. Das Ziel der Schätzung ist in jedem Fall, bestmögliche Schätzwerte für die Item- und Personenparameter zu erhalten. Bei der Schätzung der Itemschwierigkeit wird der Nullpunkt willkürlich geschätzt. Dabei bestehen grundsätzlich zwei Möglichkeiten. Bei der ersten Möglichkeit wird einem willkürlichen Item der Wert Null zugewiesen und die anderen Items werden relativ dazu geschätzt. Die zweite Möglichkeit besteht in der sogenannten Summe-Null-Normierung. Dabei ist die Summe aller Itemschwierigkeitsparameter gleich Null. Je nach Statistikprogramm ist die eine oder andere Variante vorprogrammiert (Koller et al., 2012). Daneben wird die Personenfähigkeit geschätzt. Diese wird mit dem Weighted Likelihood Estimator (WLE) nach Warm (1989) angegeben. Dabei werden die Rohwerte eines Tests nicht linear in sogenannte Logit-Einheiten (WLE-Werte) transformiert und in einer Logitskala mit Intervallqualität abgebildet. Personen mit einer grösseren Fähigkeit als die Itemschwierigkeit haben eine mehr als 50-prozentige Wahrscheinlichkeit, das Item richtig zu lösen. Hingegen werden Personen, deren Fähigkeit geringer ist als die Itemschwierigkeit, das Item mit einer geringeren Wahrscheinlichkeit als 50 % lösen. Werden Schüler- und Schülerinnen-Leistungen auf Basis des Rasch-Modells ausgewertet, können die für jedes Kind geschätzten Fähigkeitsparameter als Messwerte für die Leistung verwendet werden.

Zwei Verfahren zur Schätzung der Personenfähigkeitsparameter in Längsschnittstudien

Bei längsschnittlichen Untersuchungen empfiehlt es sich, die Personenparameter nicht zu jedem Testzeitpunkt einzeln zu berechnen, da so die Veränderung der Personenfähigkeit nicht auf einer gemeinsamen Skala abgebildet wird. Es ist hingegen sinnvoller, die dreidimensionale Datenstruktur zu berücksichtigen. Die dreidimensionale Datenstruktur bedeutet, dass die Lösungswahrscheinlichkeit einer Aufgabe in Abhängigkeit von der Personenfähigkeit, der Itemschwierigkeit und dem Zeitpunkt der Bearbeitung betrachtet wird (Hartig & Kühnbach, 2006). Wird die dreidimensionale Datenstruktur in eine zweidimensionale überführt, können Differenzwerte zwischen den messzeitpunktspezifischen Variablen gebildet und Aussagen über den individuellen Lernzuwachs der Kinder getroffen werden. Dabei können grundsätzlich zwei verschiedene Verfahren zum Einsatz kommen: die Schätzung mit virtuellen Personen und die Schätzung mit mehrdimensionalen Modellen. Bei der Schätzung mit virtuellen Personen werden die zeitpunktspezifischen Datenmatrizen vertikal an die Datenmatrix des ersten Testzeitpunkts angehängt (Rost, 2004). Dabei entstehen virtuelle Personen. Die Messungen derselben Personen zu einem späteren Zeitpunkt werden behandelt, als handele es sich um zusätzliche Personen, die das Item zum gleichen Messzeitpunkt bearbeitet hätten (Hartig & Kühnbach, 2006). Der so gebildete Datensatz wird mittels Rasch-Modell analysiert und die Parameter werden geschätzt. Durch dieses Verfahren erfolgt automatisch eine Gleichsetzung der Itemschwierigkeiten über die Messzeitpunkte. Die so ermittelten Personenparameter können auf derselben Skala interpretiert werden. Für Messwiederholungsanalysen müssen die Personenparameter wieder in die ursprüngliche Struktur überführt werden. Jede Person steht in einer Zeile und hat zu jedem Testzeitpunkt einen anderen Fähigkeitsparameter. Ein Nachteil des Verfahrens besteht indes darin, dass Abhängigkeiten zwischen den Messwerten einer Person nicht berücksichtigt werden.

Beim zweiten Verfahren wird angenommen, dass die Lösungswahrscheinlichkeit zu jedem Zeitpunkt auf eine jeweils andere Personenfähigkeitsvariable zurückgeht (Hartig & Kühnbach, 2006). Daten werden hierfür in der üblichen Struktur belassen. Damit die geschätzten Fähigkeitsparameter als Werte auf einer gemeinsamen Skala interpretiert werden können, müssen die Itemschwierigkeiten zwischen den Testzeitpunkten gleichgesetzt werden. Durch diese Restriktion im Modell wird festgelegt, dass der Schwierigkeitsparameter für jedes einzelne Item zu jedem Zeitpunkt jeweils denselben Wert annimmt.

In der Praxis zeigt sich, dass beide Verfahren annähernd zu gleichen Resultaten führen und die Personenparameterwerte praktisch identisch sind (Hartig & Kühnbach, 2006). Entscheidend ist, dass die Schwierigkeitsparameter der zwischen den Messzeitpunkten identischen Aufgaben (Ankeritems) gleichgesetzt werden, damit die Personenfähigkeit auf einer Skala abgebildet werden kann. Damit ist es auch möglich, Differenzen zwischen den einzelnen Testzeitpunkten zu bilden und als abhängige Variablen in die Analysen aufzunehmen (Rost, 2004).

6.2.3 Modellüberprüfung

Zur Überprüfung, ob für die erhobenen Daten das Rasch-Modell gilt und damit die in Abschnitt 6.2.1 beschriebenen Annahmen zutreffen, stehen verschiedene Modelltests zur Verfügung. Im Vordergrund stehen dabei Tests zur Überprüfung der Itemhomogenität und der Personenhomogenität (Bond & Fox, 2015). In längsschnittlichen Designs wird zusätzlich der Messinvarianz über die Zeit eine grosse Bedeutung zugeschrieben (Putnick & Bornstein, 2016).

Itemhomogenität

Mit dem Rasch-Modell können mittels Fit-Statistik Items identifiziert werden, die nicht oder nicht gut ins Modell passen. Die Homogenität wird dabei mittels von Abweichungsmassen für einzelne Items geprüft. Zentral sind dabei zwei itembezogene Kennwerte: die Mean-Square-Abweichungen des Typs ungewichteter (Outfit-MNSQ) und gewichteter (Infit-MNSQ) Fit-Werte. Es handelt sich dabei um ein Mass für den Grad der Verzerrung im Messmodell, wobei ein Wert von 1 Unverzerrtheit und damit eine ideale Passung in das Messmodell repräsentiert. Werte über 1 indizieren für das jeweilige Item einen Mangel an Vorhersagbarkeit der Antworten aus dem Personenschätzwert, Werte unter 1 eine zu grosse Vorhersagbarkeit. Für die Entscheidung über die Eignung von Items sind die gewichteten Abweichungen (Infit-MNSQ) von grösserer Bedeutung als die ungewichteten. In der Literatur wird kontrovers diskutiert, welche Abweichungen vom Erwartungswert tolerierbar sind (z. B. Bond & Fox, 2015). Generell werden Werte kleiner als 1 als unproblematisch angesehen, da es sich hierbei um einen Overfit handelt. Entsprechende Items passen zu gut ins Modell. Problematischer sind MNSQ-Werte grösser als 1. In diesem Fall hängt die Itemantwort zu wenig von der Eigenschaftsausprägung ab (Rost, 2004). Zur Toleranzgrenze der MNSQ-Werte werden verschiedene Richtlinien vorgeschlagen. Beispielsweise geben Boone, Staver und Yale (2014) einen weiten Bereich von 0.50 bis 1.50 an. Smith und Smith (2004) erachten indes eine Spanne zwischen 0.70 und 1.30 als akzeptabel. Bond und Fox (2015) führen für Multiple-Choice-Tests einen Bereich zwischen 0.80 bis 1.20 auf und Wilson (2005) einen solchen zwischen 0.75 und 1.33.

Personenhomogenität

Die Annahme der Personenhomogenität ist dann erfüllt, wenn alle getesteten Personen die Items aufgrund derselben Fähigkeit lösen bzw. nicht lösen können (Bond & Fox, 2015). Diese sogenannte Subgruppeninvarianz ist ein relevanter Aspekt der spezifischen Objektivität (vgl. Abschnitt 6.2.1) (Koller et al., 2012). Wenn gleichzeitig Hypothesen vorhanden sind, die den Vergleich von Gruppen betreffen, erweist sich die Überprüfung der Subgruppeninvarianz als zentral (z. B. Kinder in der Schweiz und Deutschland oder Mädchen und Jungen). Die Items müssen für die untersuchten Gruppen eine vergleichbare Schwierigkeit aufweisen. Oder anders formuliert, Personen mit gleicher Personenfähigkeit aus unterschiedlichen Gruppen müssen die gleiche Wahrscheinlichkeit haben, ein Item zu lösen. Trifft dies nicht zu, wird mindestens eine Gruppe von Personen im Test benachteiligt. Zur Überprüfung der Subgruppenvarianz werden DIF-Analysen (DIF = Differential Item Functioning) auf Itemebene vorgenommen (Schwab & Helm, 2015). Diese nehmen die Lage der Schwierigkeitsparameter in den Teilstichproben in den Blick. Die relative Schwierigkeit aller Aufgaben in Logit-Werten sollte sich über relevante Teilstichproben hinweg im Idealfall vollständig parallel darstellen bzw. die aufgabenbezogene Differenz zwischen den Teilstichproben sollte genau der Differenz der mittleren Aufgabenschwierigkeit für den ganzen Test entsprechen. Itembezogene Abweichungen der Parameter werden ausgehend vom Ideal einer Nulldifferenz berechnet. Nach Paek und Wilson (2011) können Differenzen bis maximal 0,426 bei nicht gegebener Signifikanz ($p \geq 0.05$) als vernachlässigbar gelten. Ist $p < 0.05$ und die Differenz der Parameter kleiner als 0,638, gilt der Grad an DIF als *slight to moderate*. Ist die doppelte Abweichung grösser als 0,638 (bei $p < 0.05$), gilt diese als gross. Tabelle 1 zeigt die DIF-Kategorien nach Zwick, Thayer und Levis (1999). Die hier enthaltenen Grenzwerte sind in Delta festgehalten. 1 Einheit Delta entspricht 0.426 Logits, 1.5 Einheiten Delta entsprechen 0.639 Logits.

Tabelle 1: ETS DIF-Kategorien nach Zwick, Thayer und Levis, 1999

ETS DIR Category	P(MH_CHISQ) ≤ 0.05	P(MH_CHISQ) > 0.05
ǀMH_DIFǀ ≥ 1.5	Moderate to Large	Negligible
1< ǀMH_DIFǀ < 1.5	Slight to Moderarte	Negligible
ǀMH_DIFǀ ≤ 1	Negligible	Negligible

Messinvarianz über die Zeit

Bei Längsschnittanalysen wird der Messinvarianzprüfung über die Testzeitpunkte eine grosse Bedeutung zugeschrieben (Putnick & Bornstein, 2016). Werden gleiche Items (Ankeritems) zu den verschiedenen Messzeitpunkten eingesetzt, gilt es sicherzustellen, dass diese zu den verschiedenen Messzeitpunkten das gleiche Konstrukt messen. Nur wenn dies der Fall ist, können Veränderungen in den Messwerten auf die Veränderung der mit dem Test erfassten Kompetenz zurückgeführt werden (Browne, 1984; Widaman, Ferrer & Conger, 2010). Messinvarianz über die Zeit bildet also eine zentrale Voraussetzung für längsschnittliche Analysen. Zur Überprüfung der Messinvarianz können, wie auch bei anderen Gruppenanalysen, DIF-Analysen (DIF = Differential Item Functioning) auf Itemebene durchgeführt werden (Schwab & Helm, 2015). Die Messzeitpunkte stellen dabei die Gruppen dar. Die Itemparameter werden so in einem simultanen Modell zu jedem Testzeitpunkt einzeln geschätzt und miteinander verglichen. DIF bedeutet in diesem Fall, dass die Lösungswahrscheinlichkeiten nicht nur durch die Personenfähigkeit und die Aufgabenschwierigkeit erklärt werden können, sondern dass auch der Testzeitpunkt einen Einfluss hat. Für das Vorliegen von zeitbezogener Messinvarianz gelten dieselben Grenzwerte (Zwick, Thayer & Lewis, 1999) wie bei der Messinvarianz von Subgruppen (vgl. Tabelle 1).

6.3 Erhebungsinstrumente zur Messung mathematischer Kompetenzen

Insgesamt gibt es im deutschsprachigen Raum nur wenige normierte Tests zur Messung mathematischer Kompetenzen im Kindergartenalter. Im Folgenden werden vier Tests vorgestellt und diskutiert: der MARKO-D (Ricken, Fritz & Balzer, 2013), der MBK 0 (Krajewski, 2018), der ZAREKI-K (von Aster et al., 2009) und der TEDI-MATH (Kaufmann et al., 2009).

6.3.1 MARKO-D

Beim MARKO-D (Mathematik- und Rechenkonzepte im Vorschulalter - Diagnose) von Ricken et al. (2013) handelt es sich um ein standardisiertes und normiertes Testverfahren für Kinder im Alter von vier bis sechseinhalb Jahren. Der Test wird mit jedem Kind einzeln durchgeführt und besteht aus 55 Items. Diese basieren auf dem Entwicklungsmodell von Fritz und Ricken (2008) (vgl. Abschnitt 4.2.2) und enthalten Aufgaben zu den fünf Niveaustufen Zählzahl, ordinaler Zahlenstrahl, Kardinalität und Zerlegbarkeit, Enthaltensein und Klasseninklusion sowie Relationalität. Beim Niveau I (Zählzahl) gibt es Aufgaben zum Zählen, zum Abzählen, zum Vergleichen von Mengen oder zur Erzeugung von gleichmächtigen Mengen. Niveau II (Ordinaler Zahlenstrahl) enthält Aufgaben zum Bestimmen des Nachfolgers bzw. des Vorgängers einer Zahl, zum Wegnehmen bzw. Dazulegen einer Anzahl oder zum Zusammenfügen von sichtbaren und unsichtbaren Teilmengen. Niveau III (Kardinalität und Zerlegbarkeit) umfasst Aufgaben zum Zusammenfügen oder Vermindern von Mengen, zum Weiterzählen um einen Schritt, zur Mengenseriation oder zum Erkennen von Zusammenhängen von Teilmengen und Gesamtmengen. Beim Niveau IV (Enthaltensein und Klasseninklusion) sind Aufgaben zur Bildung von zwei Teilmengen aus einer Gesamtmenge, zum Zählen in Schritten oder zum Ergänzen enthalten. Beim Niveau V (Relationalität) finden sich Aufgaben zum Bestimmen der Differenz zwischen zwei Mengen oder zwischen Zahlen oder zum Rückwärtszählen in Schritten. Die Aufgaben sind in einer Rahmengeschichte über die beiden Eichhörnchen Ben und Lisa eingebettet. Die Bearbeitungsdauer liegt zwischen 20 und 30 Minuten. Der Test wurde an einer Stichprobe von 1095 Kindern im Alter zwischen 48 und 87 Monaten normiert. Normwerte liegen für jedes Halbjahr vor. Die Gruppengrössen variieren zwischen 26 Kindern (6;6 Jahre und älter) und 329 (zwischen 5;6 und 5;11 Jahre). Der Test wurde mit dem Rasch-Modell auf Itemhomogenität (vgl. Abschnitt 6.2.3) überprüft. Alle Items hatten einen MNSQ-Wert zwischen 0.7 und 1.3 und lagen somit im tolerierten Bereich. Der Test erwies sich als objektiv durchführbar und valide. Die Reliabilität betrug .91 und .89 (Retest).

Hildenbrand (2016) setzte in ihrer Untersuchung eine Kurzform des MARKO-D längsschnittlich zu drei Testzeitpunkten ein. Auch diese Kurzversion stellte sich als verlässliches Instrument heraus. Ein Item wurde aus den Analysen ausgeschlossen, da es in der Skalierung eine von der Testnormierung abweichende Platzierung auf der Schwierigkeits- bzw. Fähigkeitsskala einnahm. Die verbliebenen 19 Items verteilten sich bei einer eindimensionalen Raschskalierung analog der Langversion. Die MNSQ-Werte lagen bei allen Items innerhalb des tolerierten Bereichs (vgl. Abschnitt 6.2.3) zwischen 0.8 und 1.26. Die Personenfähigkeit über die drei Testzeitpunkte wurde mittels virtueller Personen geschätzt, sodass die Itemschwierigkeit für alle Testzeitpunkte gleich war und die Werte auf einer Skala interpretiert werden konnten (vgl. Abschnitt 6.2.2). Deckeneffekte wurden weitgehend ausgeschlossen. „Zum ersten Erhebungszeitpunkt konnten lediglich 7 % der Kinder mehr als drei Viertel der Aufgaben richtig bearbeiten. Zum zweiten Erhebungszeitpunkte lag dieser Anteil bei 12 % und beim dritten Erhebungszeitpunkt bei 22 %" (Hildenbrand 2016, S. 166). Die Autorin schloss daraus, dass der Test auch im oberen Leistungsbereich genügend differenziert.

6.3.2 MBK 0

Der MBK 0 (Test mathematischer Basiskompetenzen im Kindergartenalter) von Krajewski (2018) ist ein Einzeltest für Kinder im Alter von 3.6 bis 7 Jahren. Diesem Test liegt das Entwicklungsmodell der Zahl-Grössen-Verknüpfung (Krajewski, 2008) zugrunde (vgl. Abschnitt 4.2.1). Die Aufgaben beziehen sich auf die im Modell beschriebenen drei Ebenen: Basisfertigkeiten, einfaches und tiefes Zahlverständnis. Es gibt eine Kurz- und eine Langform. Die Langform besteht aus acht Subtests mit insgesamt 58 Aufgaben. In der Kurzform werden fünf dieser Subtests mit 46 Aufgaben erhoben, die sich nur auf die Ebenen 1 und 2 beziehen. Die Aufgaben der ersten Ebene beziehen sich auf Kenntnisse der Zahlenfolge (Zählen vorwärts und rückwärts, Vorgänger und Nachfolger bestimmen) und der Ziffern. Auf der zweiten Ebene werden das Zuordnen von Zahlen zu Mengen, der Zahlvergleich, das Einordnen von Zahlen auf dem Zahlenstrahl und der Mengenvergleich (nur in der Langform) überprüft. Auf der dritten Ebene sollen die Kinder vorwiegend Anzahlunterschiede in verschiedenen Formen bestimmen und Rechengeschichten lösen (nur in Langform). Die Bearbeitungsdauer für die Langversion beträgt 25 Minuten, für die Kurzversion 15 Minuten. Die Normstichprobe umfasste insgesamt 3871 Kindergartenkinder. Die querschnittlich erfassten Normwerte beziehen sich auf acht Alterskategorien im Abstand von jeweils einem halben Jahr. Die interne Konsistenz liegt für die Lang- und Kurzform bei Cronbachs Alpha .95 und erwies sich entsprechend als hoch. Auch die Retest-Reliabilität nach vier bis sechs Monaten ist gut; sie betrug für beide Versionen .83 bzw. .88. Die Trennschärfen liegen für fast alle Items zwischen .40 und .70 und bewegen sich damit im mittleren Bereich. Auch die Validität ist in allen gemessenen Bereichen zufriedenstellend. Insbesondere zeigen die Befunde, dass die mathematischen Basiskomponenten, wie sie mit dem MBK 0 erfasst werden, eine gute prognostische

Aussagekraft für die mathematischen Leistungen in der Schule haben. Die Korrelationen mit den Mathematikleistungen am Ende der ersten Klasse betrugen zwischen r=.53 und r=.71. Auch mit den Mathematikleistungen am Ende der vierten Klasse gab es noch Zusammenhänge in der Grösse von r=.50.

6.3.3 ZAREKI-K

Der ZAREKI-K von von Aster et al. (2009) erfasst verschiedene Komponenten von Fähigkeiten der Zahlen- und Mengenverarbeitung und des Rechnens im Vorschulalter. Der Test wurde im Rahmen eines Projektes zur Früherkennung von Rechenstörungen entwickelt und basiert auf dem neuro-kognitiven Entwicklungsmodell des Rechnens von von Aster und Shalev (2007) (vgl. Abschnitt 4.2). Der Test wird mit jedem Kind einzeln durchgeführt und dabei mit Materialen und Bildern unterstützt. Die Durchführungsdauer liegt zwischen 25 und 40 Minuten. Der Test umfasst insgesamt 18 Aufgaben. Inhaltlich lassen sich diese Items zu drei Skalen zusammenfassen: Zählen und Zahlenwissen, Numerisches Bedeutungswissen und Rechnen sowie Arbeitsgedächtnis. Im Bereich Zählen und Zahlenwissen gibt es Items zum Zählen vorwärts, rückwärts und in Zweierschritten, zum Bestimmen des Vorgängers bzw. des Nachfolgers einer Zahl, zum Abzählen, zum Zahlen lesen und schreiben, zum Zahlen vergleichen und zur Zuordnung von Zahlen zu Mengen. Im Bereich Numerisches Bedeutungswissen und Rechnen finden sich Aufgaben zum Schätzen, zur Mengeninvarianz, zur Beurteilung von Mengen (viel, wenig), zum Rechnen und zum Zahlenstrahl. Im Bereich Arbeitsgedächtnis existieren Items zum Zahlen nachsprechen und zur Lösung von mündlich gestellten Textaufgaben. Die Normierungsstichprobe bestand aus 429 Kindern. Die Alterspanne für die meisten Kinder lag zwischen fünf und sieben Jahren. Da viele Testleistungen mit dem Alter korrelierten, wurde vor der Festlegung der Referenzgruppe überprüft, ob sich zwischen Kindern der verschiedenen Altersgruppen (60 bis 71 Monate, 72 und 83 Monate, 84 Monate und älter) Unterschiede in den Testwerten ergeben. Die Reliabilität (Cronbachs Alpha) der Skala *Zählen und Zahlenwissen* betrug .92, die der Skala *Numerisches Bedeutungswissen und Rechnen* .83 und die der Skala *Arbeitsgedächtnis* .73. Die Reliabilität des Gesamttests war mit .94 hoch. Für die Überprüfung der homogenen Leistungsanforderungen wurden Analysen mit dem Rasch-Modell durchgeführt (vgl. Abschnitt 6.2.1). Für diese Analysen wurden fünf Skalen gebildet. Für die Skalen Mengenbewusstsein, Rechnen und Zahlenstrahl waren die Ergebnisse mehrheitlich befriedigend. Für die Skala Zählfertigkeiten wurden hingegen unterschiedliche Ergebnisse festgestellt. Es konnte nicht in allen Subtests Homogenität nachgewiesen werden, teils auch aufgrund einer zu geringen Anzahl an Items. Auch in der Skala Zahlenwissen stellten sich die Analysen komplex dar. Die einzelnen Subtests (Zahlen lesen, Zahlen schreiben sowie Zahlenvergleich) waren für sich genommen in ihren Leistungsanforderungen zwar homogen, bildeten aber keine gemeinsame Skala. „Eine mögliche Erklärung liegt darin, dass die individuellen Erfahrungen im Umgang mit den Zahlen hier eine besondere Rolle spielen, und dass diese Erfahrung in den verschiedenen Bereichen sehr unterschiedlich sein kann" (von Aster et al. 2009, S. 21). Insgesamt kommen die Autoren aber zu dem Schluss, dass sich die Leistungsanforderungen in den einzelnen Subtests als homogen erwiesen und mit dem theoretischen Konzept des Erwerbs numerischer Fertigkeiten übereinstimmen. Die Überprüfung der prognostischen Validität ergab, dass 68.5 % der Kinder, die zwei Jahre nach der Testung eine Rechenstörung aufwiesen, im Kindergarten identifiziert werden konnten.

6.3.4 TEDI-MATH

Der TEDI-MATH von Kaufmann et al. (2009) erfasst die numerisch-rechnerischen Fertigkeiten vom Kindergarten bis zur dritten Klasse. Der Test differenziert insbesondere im mittleren und unteren Leistungsbereich und dient somit auch der Früherkennung von Rechenschwierigkeiten. Der Test wird als Einzelinterview mit Material- und Bildunterstützung durchgeführt. Für jede Altersgruppe gibt es eine Kern- und eine Gesamtbatterie, die sich in der Anzahl der Aufgaben voneinander unterscheiden. Die Bearbeitungsdauer der Gesamtbatterie beträgt für Kindergartenkinder durchschnittlich 30 Minuten. Die Kernbatterien für den Kindergarten umfassen zwischen fünf und acht Subtests mit insgesamt 52 bis 88 Items. Es gibt Aufgaben zu den Zählprinzipien, zum Abzählen, zur Zahlwortkenntnis, zum Rechnen mit Objektabbildungen, zur additiven Zerlegung und Textaufgaben. Die Gesamtbatterie wird durch Aufgaben zum Grössenvergleich von Zahlen, zum Ordnen nach numerischer Grösse, zur Mengeninvarianz, zur numerischen Inklusion, zur unvollständigen Addition und Subtraktion, zu den Kenntnissen arithmetischer Konzepte und zum approximativen Punktemengenvergleich ergänzt. Für die Klassen 1 bis 3 sind noch weitere Aufgaben zum Stellenwertverständnis, zum Schreiben von Zahlen, zur Addition, Subtraktion und zur Multiplikation enthalten. Die Aufgaben orientieren sich an kognitionspsychologischen und neurowissenschaftlichen Verarbeitungsmodellen der Zahlenverarbeitung und des Rechnens. In einer Untersuchung wurde zudem festgestellt, dass der Test auch die in den beiden Entwicklungsmodellen von Fritz und Ricken (2008) bzw. Krajewski (2008) postulierten numerischen Kernkompetenzen aufgreift (Fischer, Roesch & Moeller, 2017). Die Normierung des Tests erfolgte im Querschnitt pro Halbjahresklassenstufe bei insgesamt 873 deutschsprachigen Kindern. Es wurden acht verschiedene Altersgruppen zu je mindestens 100 Kindern getestet. Die Altersgruppen verteilten sich in Halbjahresschritten vom vorletzten Kindergartenjahr bis zum Ende des ersten Halbjahres der dritten Schulstufe. Es liegen je Halbjahresklassenstufe zufriedenstellende bis gute interne Konsistenz-Koeffizienten vor. Für die jeweilige Kernbatterie der verschiedenen Stufen bewegt sich Cronbachs Alpha zwischen 0.58 und 0.89. Überprüft wurde auch die Split-Half-Reliabilität, die bis auf drei Ausnahmen zwischen 0.90 und 0.96 lag und folglich als hoch bewertet werden kann. Die Ausnahmen fanden sich bei den Konsistenzkoeffizienten in der ersten Hälfte des letzten Kindergartenjahres sowie im Rechnen am Ende der zweiten und am Anfang der dritten Klasse. Die Retest-Reliabilität fiel im Gesamttest für die erste bis dritte Klasse zufriedenstellend aus, für den Kindergarten war die Retest-Reliabilität gering, was die Autorinnen und Autoren mit heterogenen Entwicklungsverläufen dieser Altersgruppe erklären. Auch die Kriteriumsvalidität, gemessen an der Korrelation zwischen Testergebnissen und Einschätzung der Erzieherinnen und Erzieher, fiel im Kindergarten sehr gering aus. Ab der Einschulung fanden sich mittelgradige bis hohe Korrelationen.

6.3.5 Diskussion der vier Testinstrumente

Alle vier vorgestellten Tests messen die mathematischen Kompetenzen von Kindergartenkindern. Der TEDI-MATH (Kaufmann et al. 2009) erfasst zusätzlich die mathematischen Kompetenzen von Kindern in der ersten bis zur Hälfte der dritten Klasse. Die Tests hatten für den Gesamttest hohe Reliabilitätswerte. Eine Ausnahme stellt jedoch der TEDI-MATH (Kaufmann et al., 2009) dar. Hier zeigten sich für die drei Kindergartenstufen der Kernbatterie nur akzeptable Werte (Cronbachs Alpha zwischen 0.58 und 0.78). Das Autorenteam erklärt diese eher niedrigen Werten durch „nicht einheitliche Vorschulprogramme im mathematischen Bereich und die stärkere Heterogenität der Entwicklungsverläufe im Kindergartenalter“ (Kaufmann et al., 2009, S. 115). Dem gegenüber stehen hohe Reliabilitätswerte für den Kindergarten der anderen Tests. Zwei der Tests wurden zusätzlich mit dem Rasch-Modell (vgl. Abschnitt 6.2.1) auf Homogenität überprüft. Beim ZAREKI-K (von Aster et al. 2009) wurde die Homogenität der Leistungsanforderungen als akzeptabel

beurteilt. Beim MARKO-D (Ricken et al., 2013) wurde die Itemhomogenität mittels der MNSQ-Werte überprüft. Alle Items wiesen Werte innerhalb der akzeptierten Grenzen auf.
Spezifische Analysen zur Personenhomogenität (Subgruppeninvarianz) (vgl. Abschnitt 6.2.3) werden in keinem Testmanual beschrieben. Der ZAREKI-K berichtet im Zusammenhang mit Rasch-Analysen zwar von Teilungskriterien in Bezug auf Alter, Geschlecht und Testleistung. „Als Teilungskriterium wurde das Alter der Kinder herangezogen mit einer Gruppe jüngerer Kinder [...] und einer Gruppe von älteren Kindern [...]. Die Stabilität der Ergebnisse wurde mit zwei weiteren Kriterien für die Gruppeneinteilung, Geschlecht und Testleistung, überprüft“ (von Aster et al. 2009, S. 21). Es werden aber keine weiteren Ergebnisse dieser Analysen berichtet. Im TEDI-MATH wurden die Mittelwerte zwischen Jungen und Mädchen mittels t-Test analysiert und bei signifikanter Abweichung die geschlechtskorrigierten Rohwerte zur Berechnung der Normwerte verwendet. Beim MBK 0 wurde auf separate Normwerte für Jungen und Mädchen verzichtet, „obwohl hier teilweise signifikante Vorteile der Jungen gefunden wurden (mit Effektstärken zwischen d=.05 und d=.41)“ (Krajewski, 2018, S. 32).

Im Zusammenhang mit der Messung mathematischer Kompetenzen im Längsschnitt muss festgehalten werden, dass die hier beschriebenen Tests mehrheitlich andere Zielsetzungen priorisieren. Mit den Tests sollen die mathematischen Kompetenzen im Vorschulbereich erfasst und mit einer Normstichprobe verglichen werden. Damit können statusdiagnostische Aussagen über den mathematischen Entwicklungsstand eines Kindes gemacht werden. Der ZAREKI-K beispielsweise wurde als Screenings-Instrument für eine Risikoeinschätzung späterer Rechenstörungen entwickelt. Mit der Früherkennung von mathematischen Defiziten wird eine „gezielte und individuell adaptierte Förderung“ (von Aster et al. 2009, S. 3) möglich. Der MARKO-D spricht von einer entwicklungsorientierten Diagnostik, die es ermöglicht, „den individuellen Entwicklungsstand eines Kindes zu beschreiben und einer Entwicklungssequenz zuzuordnen“ (Ricken et al. 2013, S. 5). Der TEDI-MATH dient „vor allem der Diagnosestellung einer eventuell vorliegenden Dyskalkulie sowie der Interventionsplanung für eine individuell auf das Leistungsprofil des Kindes abgestimmten Dyskalkulie-Therapie“ (Kaufmann et al., 2009, S. 11). Entsprechend wurden die Tests querschnittlich normiert. Die Daten der Normierung für verschiedene Altersgruppen stammen von unterschiedlichen Stichproben, die so nicht verglichen werden können. Der MBK O kann sowohl für die diagnostische Abklärung als auch in der Forschung eingesetzt werden; hier „zur differenzierten Frühdiagnostik der mathematischen Kompetenzentwicklung sowie zur Evaluation von Interventionsmassnahmen“ (Krajewski, 2018, S. 8). Über die Eignung des Instruments zur längsschnittlichen Messung, beispielsweise über zeitbezogene Messinvarianz oder über mögliche Boden- und Deckeneffekte, liegen noch keine Ergebnisse vor. Der MAKRO-D (Ricken et al. 2013) wurde in der Studie von Hildenbrand (2016) zwar längsschnittlich eingesetzt, es wurden aber keine Angaben zur Überprüfung der Messinvarianz über die drei Testzeitpunkte gemacht, ein Desiderat für längsschnittliche Messungen (Putnick & Bornstein, 2016).

In der hier vorliegenden Untersuchung wurde der TEDI-MATH von Kaufmann et al. (2009) zur längsschnittlichen Erfassung im Kindergarten eingesetzt. Der TEDI-MATH besitzt gegenüber den anderen Tests den Vorteil, dass er eine grössere Alterspanne umfasst. Um einen Deckeneffekt zu vermeiden und der grossen Heterogenität sowie der rasch verlaufenden Entwicklung von Kindern zwischen vier und sechs Jahren gerecht zu werden, wurden deshalb auch Aufgaben eingesetzt, die für ältere Kinder vorgesehen wurden. Zudem wurden einige Anpassungen vorgenommen und zwei Subtests neu entwickelt (vgl. Abschnitt 7.3.1). Der Test wurde mit dem Rasch-Modell zur Eignung der längsschnittlichen Messung überprüft (vgl. Abschnitt 8.1).

6.4 Zusammenfassung

In der klassischen Testtheorie (KTT) wird das Ergebnis eines Tests als ein mit Messfehlern behaftetes Merkmal verstanden. Das Ziel der KTT besteht darin, die Messfehler möglichst klein zu halten. Dabei sind drei zentrale Gütekriterien von Bedeutung: Objektivität, Reliabilität und Validität. Unter Objektivität wird verstanden, dass die Testergebnisse unabhängig von der Person sind, die den Test durchführt bzw. auswertet. Reliabilität meint die Zuverlässigkeit eines Tests. Diese wird in der KTT üblicherweise mit Cronbachs Alpha angegeben. Als drittes Gütekriterium gilt die Annahme der Validität. Ein Test erfüllt dieses Kriterium, wenn er das zu messende Konstrukt möglichst genau und erschöpfend misst. In Ergänzung zur klassischen Testtheorie wird bei der Konstruktion von Leistungs- und Intelligenztests zunehmend auch das Rasch-Modell zur Prüfung des Instruments herangezogen. Dieses gehört zur Item-Response-Theorie (IRT), einer probabilistischen Testtheorie, die davon ausgeht, dass Testergebnisse Indikatoren latenter Dimensionen sind. Im Rasch-Modell gilt deshalb, dass die Wahrscheinlichkeit, ein Item zu lösen, ausschliesslich von der Fähigkeit der Person und der Itemschwierigkeit abhängt. Die Parameter der Personenfähigkeit bzw. der Itemschwierigkeit werden im Rasch-Modell geschätzt. Bei längsschnittlichen Untersuchungen kann die

Itemschwierigkeit über die drei Messzeitpunkte fixiert werden. Dadurch können die Werte auf derselben Skala interpretiert werden. So kann die Differenz der Fähigkeitswerte zwischen den Testzeitpunkten als Fähigkeitszuwachs interpretiert werden. Mit dem Rasch-Modell lässt sich das Testinstrument auf Item- bzw. Personenhomogenität überprüfen. Items, die aufgrund der Personenparameter und der Itemschwierigkeit empirisch zu stark von der erwarteten Lösungshäufigkeit abweichen, sollten aus den Analysen ausgeschlossen werden. Entscheidend für die Eignung von Items sind die gewichteten Abweichungen, die sogenannten Infit-MNSQ-Werte. Zur Überprüfung der Personenhomogenität werden DIF-Analysen durchgeführt. Damit wird überprüft, ob die Items für alle Personen einer Gruppe gleich schwierig bzw. gleich einfach sind. Ähnlich kann bei längsschnittlichen Untersuchungen auch die Messinvarianz über die Zeit überprüft werden. Dabei wird untersucht, ob die Items zu jedem Testzeitpunkt das gleiche Konstrukt messen. Items, die dieses Kriterium nicht erfüllen, sollten aus den Tests ausgeschlossen werden. Insgesamt liegen im deutschsprachigen Raum nur wenige Tests vor, die die numerischen Kompetenzen von Kindergartenkindern messen. Die vorgestellten Tests (MARKO-D, MBK 0, ZAREKI-K, TEDI-MATH) verfolgen vor allem eine diagnostische Zielsetzung. Sie sollen die mathematischen Kompetenzen von Kindergartenkindern valide messen, damit diese mit einer Normstichprobe verglichen und somit Aussagen über den Entwicklungsstand getroffen werden können. Die Eignung zur längsschnittlichen Erfassung (Messinvarianz über die Zeit) wurde nicht überprüft.

7 Darstellung der Untersuchung und des methodischen Vorgehens

Die vorliegende Untersuchung ist Teil einer vom Schweizerischen Nationalfonds und der Deutschen Forschungsgemeinschaft unterstützen Studie[1] mit dem Titel „Struktur fachspezifischer professioneller Kompetenzen von pädagogischen Fachkräften und ihre differenziellen Effekte auf die Qualität von mathematischen Lehr-Lern-Situationen im Kindergarten und den Kompetenzzuwachs von Kindern" (intern WILMA: Wir lernen Mathematik). Im Zentrum dieser Studie stehen die längsschnittliche Messung mathematischer Leistungen von Kindergartenkindern und Unterschiede in der Entwicklung dieser mathematischen Kompetenzen. Im Folgenden werden zuerst die Forschungsfragen der Studie vorgestellt. Mit Bezug zum ersten Teil dieser Arbeit (Kapitel 2 bis 6) wurden zudem Hypothesen generiert, die in diesem Zusammenhang ebenfalls präsentiert werden. Danach werden das Untersuchungsdesign, die Erhebungsinstrumente und die Stichprobe erläutert. Der letzte Abschnitt widmet sich schliesslich den Analysemethoden, die für die Beantwortung der Fragestellungen relevant sind und in dieser Untersuchung eingesetzt wurden.

7.1 Fragestellungen und Hypothesen der eigenen Studie

Mit dieser Untersuchung werden zwei Hauptziele verfolgt: Erstens die Überprüfung eines Mathematiktests zur validen Erfassung mathematischer Kompetenzen im Längsschnitt und zweitens die Analyse von Unterschieden in der mathematischen Leistungsentwicklung von Kindergartenkindern.

Das erste Ziel orientiert sich an den Ausführungen zur längsschnittlichen Erfassung mathematischer Kompetenzen in Kapitel 6. Die Messung von Kompetenzen über mehrere Testzeitpunkte stellt hohe Anforderungen an Testinstrumente. So muss das Testinstrument verschiedene Kriterien erfüllen. Der in dieser Studie verwendete Test, der TEDI-MATH von Kaufmann et al. (2009), wurde in der Originalversion querschnittlich normiert und erfüllt die Gütekriterien der klassischen Testtheorie. Es existieren aber bisher keine Analysen zur Beurteilung seiner Eignung zur Erfassung längsschnittlicher Daten. Es stellt sich deshalb die

[1] SNF-Projekt Nr. 156680

Frage, ob eine adaptierte Form des TEDI-MATH die Kriterien für eine Messung im Längsschnitt erfüllt und der Überprüfung mit dem Rasch-Modell aus der Item-Response-Theorie standhalten kann.

Das zweite Ziel betrifft die aktuelle Forschung zu frühen mathematischen Kompetenzen. So zeigen die Ausführungen in Kapitel 3 einheitlich, dass mathematische Kompetenzen von Kindergartenkindern ein wichtiger Prädiktor für Rechenleistungen in der Schule und deshalb von grosser Bedeutung sind. Unter frühen mathematischen Kompetenzen werden hier Aspekte wie Mengenerfassung, Zahlaspekte, Zählentwicklung, Teil-Ganzes-Konzept und erste zählende Rechenleistungen verstanden (vgl. Abschnitt 3.1). Sie entwickeln sich ab der frühsten Kindheit und werden in Entwicklungsmodellen systematisch dargestellt (vgl. Abschnitt 4.2). Die Entwicklung dieser Kompetenzen erfolgt jedoch nicht für alle Kinder gleich und wird von verschiedenen Faktoren beeinflusst. Der Forschungsstand zum Einfluss von individuellen und kontextuellen Faktoren auf die mathematische Entwicklung innerhalb des Kindergartens ist bescheiden, und die Befunde sind nicht immer einheitlich (vgl. Kapitel 5). Zudem fehlen Forschungsarbeiten zu unterschiedlichen Entwicklungsverläufen von Kindergartenkindern, beispielsweise von Kindern mit niedrigem, mittlerem und hohem Vorwissen. Im Hinblick auf präventive Massnahmen ist es deshalb wünschenswert, die Entwicklungsverläufe von Kindergartenkindern zu analysieren, um Einflussfaktoren auf diese vorschulische mathematische Entwicklung zu eruieren.

Ausgehend von diesen Lücken wurden vier Forschungsfragen formuliert und Hypothesen generiert. Die erste Fragestellung bezieht sich auf die Überprüfung des Testinstrumentes (Abschnitt 7.1.1). Die zweite Fragestellung fokussiert die mathematischen Kompetenzen der Kinder zum ersten Testzeitpunkt. Von Interesse ist dabei, über welche Kompetenzen die Kinder verfügen und welche Faktoren Unterschiede in den Leistungen erklären können (Abschnitt 7.1.2). Die dritte Fragestellung bezieht sich auf die Entwicklung der mathematischen Kompetenzen im Kindergarten über drei Testzeitpunkte und auf Einflussfaktoren auf diese Entwicklung (Abschnitt 7.1.3). Die vierte Fragestellung fokussiert schliesslich die Entwicklung mathematischer Kompetenzen von Kindern mit niedrigen, mittleren und hohen Kompetenzen über die drei Testzeitpunkte (Abschnitt 7.1.4).

7.1.1 Überprüfung des Mathematiktests

Diese Fragestellung nimmt Bezug auf die Eignung des Testinstruments zur Messung mathematischer Kompetenzen im Längsschnitt. Sie lautet wie folgt:

1. Erfüllt der in dieser Studie eingesetzte, adaptierte Test TEDI-MATH (Kaufmann et al., 2009) die Kriterien einer validen längsschnittlichen Messung mathematischer Kompetenzen von Kindergartenkindern?

Zur Untersuchung dieser Frage wurden neben der üblichen Prüfung der Reliabilität mit Cronbachs Alpha aus der klassischen Testtheorie mit dem Rasch-Modell der Item-Response-Theorie die einzelnen Items auf ihre Passung von erwarteter und empirisch beobachteter Lösungswahrscheinlichkeit in Abhängigkeit von Personenfähigkeit und Itemschwierigkeit untersucht. Ebenfalls mit dem Rasch-Modell wurden die einzelnen Items bezüglich Messinvarianz über die drei Testzeitpunkte und bezüglich Subgruppeninvarianz von Mädchen und Jungen und von Kindern in Deutschland und der Schweiz überprüft (vgl. Abschnitt 6.2.3).

7.1.2 Mathematische Kompetenzen von Kindern zum ersten Messzeitpunkt und ihre Einflussfaktoren

Die folgende Fragestellung bezieht sich auf das mathematische Wissen von Kindern zum ersten Testzeitpunkt. Herausgefunden werden sollte hiermit, über welche mathematischen Kompetenzen die Kinder verfügen, wie unterschiedlich diese ausgeprägt sind und welche Faktoren die Unterschiede zu erklären vermögen.

2. Über welche mathematischen Kompetenzen verfügen Kinder beim ersten Testzeitpunkt und durch welche Kontextfaktoren werden diese Kompetenzen beeinflusst?

Ausgehend von bisherigen Studien (Hasemann & Gasteiger, 2014; Moser et al., 2004) wurde auch in dieser Untersuchung eine grosse Heterogenität bezüglich früher mathematischer Kompetenzen erwartet.

Verschiedene Forschungsarbeiten haben einen Zusammenhang von Intelligenz und mathematischen Kompetenzen offengelegt (z. B. Tiedemann & Billmann-Mahecha, 2004), aber auch darüber berichtet, dass der Einfluss der Intelligenz sinkt, wenn das mathematische Vorwissen ins Modell aufgenommen wird (z. B. Krajewski & Schneider, 2009). Da zum ersten Testzeitpunkt noch keine Daten zum mathematischen Vorwissen der Kinder vorlagen, konnte davon ausgegangen werden, dass die Intelligenz, analog zur Studie von Sale et al. 2018, einen Einfluss auf die mathematischen Leistungen zu diesem Testzeitpunkt hatte.

Der Einfluss des Geschlechts auf die mathematischen Kompetenzen im Kindergarten wurde von mehreren Forschungsgruppen untersucht. Bei sechs der untersuchten Studien hatte das Geschlecht keinen Einfluss (Bonny & Lourenco, 2013; Dornheim, 2008; Niklas & Schneider, 2012; Rohe & Quaiser-Pohl, 2010; Sahr, 2012, Sale et al. 2018). In zwei Studien wiesen die Mädchen zu Beginn des Kindergartens höhere Leistungen auf (Anders et al., 2012; Weinhold Zulauf et al., 2003) und in einer Studie im Bereich des Zahlenvorwissens die Jungen (Krajewski, 2008). Insgesamt gibt es mehr Studien, die keinen Geschlechterunterschied im Kindergarten finden konnten als Studien mit Geschlechterunterschied. Deshalb wurde in dieser Untersuchung von der Hypothese ausgegangen, dass das Geschlecht keinen Einfluss auf die mathematischen Kompetenzen zum ersten Testzeitpunkt hat.

Zum Einfluss des Alters sind Hinweise vorhanden, wonach ältere Kinder im Vorschulalter höhere mathematische Kompetenzen aufweisen als jüngere (Jörns et al., 2013; Sale et al. 2018). Zudem ist bekannt, dass sich mathematische Kompetenzen seit frühster Kindheit kontinuierlich entwickeln (vgl. Kapitel 4). Es konnte deshalb davon ausgegangen werden, dass ältere Kinder beim ersten Testzeitpunkt insgesamt höheren mathematische Kompetenzen haben, da der mathematische Kompetenzerwerb um einige Wochen oder Monate länger möglich war.

Bezüglich des Einflusses von Merkmalen der sozialen Umwelt wurde in dieser Studie nur der Einfluss der Erstsprache analysiert. Da keine Daten zum sozioökonomischen Status und zum elterlichen Unterstützungsverhalten vorliegen, konnte der Einfluss dieser Faktoren nicht untersucht werden. Der Migrationshintergrund wurde zwar erfasst (vgl. Abschnitt 7.3.3), korreliert aber sehr stark mit der Erstsprache (r = .71). Bei den Analysen musste deshalb von Multikollinearität ausgegangen werden, was zu einer ungenauen Schätzung führt. Eine in Wirklichkeit bedeutsame Variable kann einen nicht signifikanten Regressionskoeffizienten aufweisen (Bortz & Schuster, 2016). Deshalb wurde in vorliegender Studie nur der Einfluss der

Erstsprache als Merkmal sozialer Umwelt untersucht. Der Forschungsstand zum Einfluss der Erstsprache auf die mathematischen Kompetenzen im Kindergarten ist unbefriedigend. Die wenigen Forschungsarbeiten zeigen tendenziell, dass Kinder mit einer anderen Erstsprache als Deutsch in Bezug auf mathematische Kompetenzen benachteiligt sind (vgl. Abschnitt 5.5.3). Es wurde deshalb die Annahme getroffen, dass auch in dieser Untersuchung die Erstsprache einen signifikanten Einfluss auf die mathematischen Kompetenzen zum ersten Testzeitpunkt hat.

Im Hinblick auf Kontextfaktoren, die nicht das Individuum betreffen, wurde in dieser Studie der Einfluss des Landes zur Erklärung von Unterschieden zwischen Kindern in Deutschland und der Schweiz untersucht. Wie in Kapitel 2 aufgearbeitet, unterscheiden sich die beiden Länder hinsichtlich verschiedener Aspekte in der frühen Bildung. Diese Unterschiede können sich allenfalls auf die Kompetenzentwicklung der Kinder während des Kindergartens auswirken. Es ist aber kaum zu erwarten, dass sich Kinder in der Schweiz und in Deutschland bereits zum ersten Testzeitpunkt in ihren mathematischen Kompetenzen unterscheiden. Zudem sind sie vom Alter her vergleichbar. Es kann deshalb von ähnlichen lern- und entwicklungspsychologischen Voraussetzungen ausgegangen werden.

Hypothesen zur zweiten Forschungsfrage:

- Die mathematischen Kompetenzen beim ersten Testzeitpunkt sind von den allgemeinen kognitiven Fähigkeiten, dem Alter und der Erstsprache abhängig. Kinder mit höheren kognitiven Leistungen und ältere Kinder haben höhere Kompetenzen als Kinder mit niedrigen kognitiven Leistungen und jüngere Kinder. Ebenso zeigen Kinder mit Erstsprache Deutsch höhere mathematische Leistungen als Kinder mit einer anderen Erstsprache.
- Das Geschlecht hat keinen Einfluss auf die mathematischen Kompetenzen zum ersten Testzeitpunkt.
- Es gibt keinen Unterschied in Bezug auf mathematische Kompetenzen zwischen Kindern in Deutschland und Kindern in der Schweiz.

7.1.3 Mathematische Kompetenzentwicklung im Kindergarten und ihre Einflussfaktoren

In Kapitel 5 wurde der Forschungsstand zum Zusammenhang von Kontextfaktoren und mathematischen Kompetenzen erarbeitet. Insgesamt liegen wenige Forschungsarbeiten vor, die den Einfluss von individuellen und kontextuellen Faktoren auf die Leistungsentwicklung innerhalb des Kindergartens untersuchten. Diese Studie leistet einen Beitrag zur Schliessung dieser Forschungslücke. Folgende Frage steht dabei im Zentrum:

3. Wie entwickeln sich mathematische Kompetenzen von Kindergartenkindern über drei Testzeitpunkte und welchen Einfluss haben individuelle und kontextuelle Faktoren auf diese Entwicklung?

Die Leistungsentwicklung über drei Messzeitpunkte im Kindergarten wurde in dieser Form noch nicht untersucht. Analog zu Forschungsarbeiten aus der Grundschule (Ditton & Krüsken, 2009; Fritz et al., 2018; Karst & Lipowsky, 2013) wurde von einer linearen Leistungsentwicklung ausgegangen und von einer reduzierten Leistungsstreuung beim zweiten und dritten Testzeitpunkt.

Verschiedenste Untersuchungen haben gezeigt, dass das mathematische Vorwissen der Kinder der bedeutendste Prädiktor für spätere mathematische Kompetenzen ist (z. B. Dornheim, 2008; Gallit et al., 2018; Manfra et al., 2017). Entsprechend wurde auch in dieser Untersuchung davon ausgegangen, dass die mathematischen Kompetenzen beim ersten Testzeitpunkt die Unterschiede in den Leistungen beim zweiten bzw. dritten Testzeitpunkt am besten zu erklären vermögen.

Zum Einfluss der Intelligenz liegen verschiedene Ergebnisse vor. In einigen Studien hat sich die Intelligenz auch unter Einbezug des Vorwissens als signifikanter Prädiktor in der Vorhersage von Rechenleistungen erwiesen. In anderen Studien leistete die Intelligenz unter Einbezug des Vorwissens hingegen keinen Beitrag zur Varianzaufklärung in den mathematischen Leistungen oder nur einen indirekten (vgl. Abschnitt 5.2). Die Studie von Hauser et al. (2014) ist bezüglich des Studiendesigns und der untersuchten Stichprobe mit dieser Untersuchung vergleichbar. Dabei liess sich ein Einfluss der Intelligenz auf die mathematischen Kompetenzen feststellen. Es wurde deshalb auch für die hier vorliegende Untersuchung vermutet, dass die kognitiven Fähigkeiten in dieser Studie auch unter Einbezug des mathematischen Vorwissens einen Einfluss auf die Kompetenzentwicklung haben.

Analog zur Studie von Anders et al. (2012) wurde angenommen, dass der Kompetenzzuwachs der Jungen gegenüber den Mädchen höher ist und dass ältere Kinder und Kinder mit Deutsch als Erstsprache einen höheren Leistungszuwachs haben.

Zur Erklärung von Unterschieden in der Kompetenzentwicklung auf Klassenebene wurde aufgrund der verschiedenen vorschulischen Bildungssysteme in Deutschland und der Schweiz auch das Land als möglicher Prädiktor in Betracht gezogen (vgl. Kapitel 2). Allerdings sah das Design dieser Studie vor, dass sowohl in Deutschland als auch in der Schweiz alle pädagogischen Fachkräfte zehn Regelspiele zur mathematischen Förderung erhalten, die sie im Kindergartenalltag zwei- bis dreimal pro Woche einsetzen (vgl. Abschnitt 7.2). Deshalb war davon auszugehen, dass sich die länderspezifischen Unterschiede nicht auf die Kompetenzentwicklung der Kinder auswirken, da durch den Einsatz der Spiele alle Kinder eine vergleichbare mathematische Förderung erhielten.

In Bezug auf den Einfluss der Ausbildung der pädagogischen Fachkräfte legen Forschungsergebnisse aus dem Kindergarten zwar nahe, dass das pädagogische Fachwissen in Mathematik ein nicht unwesentlicher Faktor für die Planung von Lernsituationen und damit zentral für das Lernen der Kinder ist (Dunekacke et al., 2015). Allerdings konnte nicht angenommen werden, dass dieses Fachwissen allein auf die Art der Ausbildung (erfasst über akademisch versus nicht akademisch, vgl. Abschnitt 7.3.3) zurückzuführen ist.

Hypothesen zur dritten Forschungsfrage:

- Die Entwicklung mathematischer Kompetenzen über die drei Testzeitpunkte ist linear und die Kompetenzstreuung nimmt vom ersten bis zum dritten Testzeitpunkt ab.
- Die mathematische Kompetenzentwicklung wird durch das Vorwissen, die kognitiven Fähigkeiten, das Geschlecht, das Alter und die Erstsprache der Kinder beeinflusst.
    - Das Vorwissen erklärt den grössten Teil der Unterschiede in den mathematischen Kompetenzen beim zweiten bzw. dritten Testzeitpunkt. Je höher das mathematische Vorwissen war, desto höher sind die Kompetenzen zum zweiten bzw. zum dritten Testzeitpunkt.

    - Kinder mit höheren kognitiven Fähigkeiten haben einen grösseren Kompetenzzuwachs als Kinder mit niedrigen kognitiven Fähigkeiten.
    - Jungen haben einen grösseren Kompetenzzuwachs als Mädchen.
    - Ältere Kinder haben einen grösseren Kompetenzzuwachs als jüngere Kinder.
    - Kinder mit Erstsprache Deutsch haben einen grösseren Kompetenzzuwachs als Kinder mit einer anderen Erstsprache.
- Kinder in der Schweiz und in Deutschland unterscheiden sich nicht in Bezug auf ihren mathematischen Kompetenzzuwachs.
- Die Ausbildung der Fachkräfte hat keinen Einfluss auf die mathematische Kompetenzentwicklung der Kindergartenkinder.

7.1.4 Entwicklungsverlauf von Kindern mit unterschiedlichem Vorwissen

Die letzte Forschungsfrage widmet sich dem Entwicklungsverlauf von Kindern mit unterschiedlichem Vorwissen:

4. Wie entwickeln sich die mathematischen Kompetenzen von Kindern mit niedrigem, mittlerem und hohem Vorwissen und welchen Einfluss haben Kontextfaktoren auf diese Entwicklung?

Hinsichtlich der Kompetenzentwicklung von Kindern mit unterschiedlichem Vorwissen gibt es Hinweise aus der Grundschule, wonach die mathematische Entwicklung nicht parallel verläuft. Kinder mit niedrigem Vorwissen weisen den grössten Kompetenzzuwachs auf, während Kinder mit hohem Vorwissen den geringsten Leistungszuwachs erzielen (Ditton & Krüsken, 2009; Fritz et al., 2018; Karst & Lipowsky, 2013). Es konnte folglich davon ausgegangen werden, dass die Entwicklung von Leistungsgruppen im Kindergarten ähnlich verläuft. Zum Einfluss von Kontextfaktoren in verschiedenen Leistungsgruppen liegen keine Forschungsergebnisse vor.

Hypothesen zur vierten Forschungsfrage:

- Kinder mit niedrigem mathematischem Vorwissen haben den höchsten mathematischen Leistungsfortschritt, gefolgt von Kindern mit mittlerem

Vorwissen. Kinder mit hohem Vorwissen haben den kleinsten Kompetenzzuwachs. Entsprechend reduziert sich die Leistungsstreuung beim zweiten und beim dritten Testzeitpunkt.

7.2 Untersuchungsdesign

Mit vorliegender Untersuchung sollen die mathematischen Kompetenzen von Kindergartenkindern und deren Entwicklung analysiert werden. Zusätzlich soll der Einfluss von individuellen und kontextuellen Faktoren auf die Kompetenzen und deren Entwicklung überprüft werden. Das Design der Studie ist daher so angelegt, dass diese Fragen beantwortet und die Hypothesen geprüft werden können (Tabelle 2).

Tabelle 2: Studiendesign

Testzeitpunkt 1	Testzeitpunkt 2		Testzeitpunkt 3
Mathematische Kompetenzen Alter, Geschlecht, Erstsprache, Land	Mathematische Kompetenzen Kognitive Fähigkeiten	Alle Kindergärten erhalten zehn Regelspiele zur mathematischen Förderung	Mathematische Kompetenzen

Die mathematischen Kompetenzen der Kinder wurden zu drei Testzeitpunkten erhoben. Im März 2016 wurden die Kinder zum ersten Mal mit einer adaptierten Form des TEDI-MATH von Kaufmann et al. (2009) überprüft (vgl. Abschnitt 7.3.1). Gleichzeitig wurden auch Daten wie das Alter, das Geschlecht und die Erstsprache erhoben (vgl. Abschnitt 7.3.3). Im September 2016 wurden die Kinder erneut auf ihre mathematischen Kompetenzen hin getestet. Zu diesem Zeitpunkt wurden auch die kognitiven Fähigkeiten erfasst (vgl. Abschnitt 7.3.2). Im Juni 2017, kurz vor der Einschulung, erfolgte schliesslich ein dritter Test der Kindergartenkinder mit dem TEDI-MATH. Die Erhebungen wurden von geschulten Testleiterinnen durchgeführt, die sich vorgängig bewerben mussten. Es handelte sich mehrheitlich um Studierende aus dem Fach Erziehungswissenschaft oder Psychologie, die Erfahrungen im Umgang mit jungen Kindern vorweisen konnten. In einer dreistündigen Schulung wurden sie darüber informiert, wie der Test durchzuführen ist und worauf besonders geachtet werden sollte. Die Erfassung der kognitiven Fähigkeiten haben dieselben Testleiterinnen durchgeführt. Auch hierfür wurden sie vorgängig geschult.

Die pädagogischen Fachkräfte erhielten zwischen dem zweiten und dritten Testzeitpunkt zehn Regelspiele zur mathematischen Förderung. Dies war Teil der Gesamtstudie WILMA (vgl. Kapitel 7, Einleitung), in der unter anderem der Einfluss von Fortbildungen auf den Einsatz der Spiele und der Spielbegleitung erforscht wurde. Da diese Intervention jedoch nicht Teil der vorliegenden Untersuchung war, bezieht sich auch keine der Forschungsfragen auf den Einsatz oder die Wirkung der Spiele. Die Spiele werden an dieser Stelle aber trotzdem vorgestellt, da deren Verwendung allenfalls Unterschiede in der Kompetenzentwicklung zwischen den Testzeitpunkten erklären kann. Die Brett- und Kartenspiele wurden mehrheitlich im SpiMaF-Projekt der Pädagogischen Hochschule St. Gallen entwickelt und erprobt (Hauser et al., 2014). Tabelle 3 gibt einen Überblick über die eingesetzten Spiele und die mathematischen Kompetenzen, die damit gefördert werden (vgl. Hauser, Rathgeb-Schnierer, Stebler & Vogt, 2015, S. 64).

Tabelle 3: Spiele zur mathematischen Förderung

Spiel	Beschreibung	Förderschwerpunkt
Fünferraus	Adaption des im Handel erhältlichen Elferraus	Zahlen lesen Zahlenreihenfolge kennen
Klecksimonster	Selbstentwickeltes Kartenspiel	Zuordnung von Zahlen und Mengen
Klipp-Klapp	Im Handel als Shut the Box erhältlich Angepasste Regeln	Zuordnung von Zahlen und Mengen Zusammensetzen und Zerlegen von Mengen
Mehr ist mehr	Selbstentwickeltes Kartenspiel	Vergleichen von Mengen
Stechen	Selbstentwickeltes Kartenspiel	Vergleichen von Mengen
Pasch	Im Handel als Yahtzee erhältlich Angepasste Regeln	Zuordnung von Zahlen und Mengen

Steine sammeln	Selbstentwickeltes Spiel	Bestimmen von Anzahlen
Verflixte fünf	Adaption des im Handel erhältlichen Hornochsen-Spiel	Zusammensetzen und Zerlegen von Mengen Erstes Rechnen
Goldstückspiel	Brettspiel: entwickelt für eine frühere Studie (Moser Opitz, 2008)	Zählen Bestimmen von Anzahlen Eins-zu-Eins-Zuordnung
Geheimniskrämerei	Entwickelt innerhalb der WILMA-Studie (vgl. Kapitel 7, Einleitung)	Teil-Ganzes-Konzept

Die Spiele können in Partnerarbeit oder in Gruppen zu dritt oder zu viert gespielt werden. Alle Spiele können zudem variiert werden, sodass sowohl Kinder mit niedrigen als auch Kinder mit hohen mathematischen Kompetenzen ihrem Niveau entsprechend spielen können. Es wurde den pädagogischen Fachkräften freigestellt, welche Spiele sie wann einführen und auch welche Spiele sie welchen Kindern geben. Empfohlen wurde einzig, die Spiele ca. zwei- bis dreimal pro Woche während ca. 30 Minuten zu spielen.

7.3 Erhebungen

Zur empirischen Erfassung von Merkmalsausprägungen, die zur Beantwortung der Fragestellung dienen sollen, wurden die mathematischen Kompetenzen mit dem TEDI-MATH (Kaufmann et al., 2009) und die kognitiven Fähigkeiten mit einer Kurzform des CFT 1-R (Weiss & Osterland, 2012) erfasst. Daneben wurden Hintergrundvariablen erhoben.

7.3.1 Mathematische Kompetenzen

Für die vorliegende Untersuchung wurde ein standardisierter Einzeltest für die Erfassung mathematischer Kompetenzen benötigt, der mit Kindern im Alter von vier bis sechs Jahren in ca. 15 bis 25 Minuten durchgeführt werden kann. Der TEDI-MATH (Kaufmann et al., 2009) hatte sich für das geplante Forschungsvorhaben als geeignet herausgestellt. Der Test basiert auf

Erkenntnissen zur Zahlbegriffsentwicklung (vgl. Abschnitt 6.3.4). Eine Normierung liegt für die Altersspanne vom vorletzten Kindergartenjahr bis zum Ende des ersten Schulhalbjahres der dritten Klasse vor. Somit ermöglicht das Instrument die Messung eines breiten Spektrums mathematischer Kompetenzen sowohl im tiefen als auch im hohen Leistungsbereich. Dies war für die vorliegende Untersuchung insofern von Bedeutung, als der Test längsschnittlich über drei Messzeitpunkte eingesetzt wurde und ein Decken- bzw. Bodeneffekt vermieden werden sollte. Für jede Altersgruppe existieren eine Kern- und eine Gesamtbatterie, die sich in der Anzahl der Aufgaben unterscheiden. Im nächsten Abschnitt folgt ein Überblick über die verwendeten Subtests, die im Anschluss erläutert werden. Allfällige Änderungen oder Anpassungen gegenüber dem TEDI-MATH original werden beschrieben und begründet.

Überblick Subtests und Items

Tabelle 6 fasst die in dieser Studie verwendeten Subtests sowie die Anzahl der Items und Punkte zusammen. Beim zweiten und dritten Testzeitpunkt wurde der Test um einige schwierige Items ergänzt.

Tabelle 4: Übersicht Subtests und Anzahl Items zu den drei Testzeitpunkten

Subtest	Anzahl Items/Anzahl Punkte	
	Testzeitpunkt 1	Testzeitpunkte 1 und 2
1. Zahlwortreihe	7/8	9/11
2. Abzählen	10/10	10/10
3. Entscheidung arabische Zahl	1/1	1/1
4. Grössenvergleich arabische Zahl	8/8	8/8
5. Transkodieren (Zahlen lesen)	15/15	15/15
6. Ordnen nach numerischer Grösse	2/2	2/2
7. Ordnen nach numerischer Grösse	2/2	2/2
8. Zahlen-Grössen-Zuordnung	1/2	1/2
9. Numerische Inklusion	3/3	3/3
10. Additive Zerlegung	2/2	4/4
11. Rechnen mit Objektabbildungen	6/6	6/6
12. Addition	9/9	9/9

13. Unvollständige Addition	4/4	4/4
14. Unvollständige Subtraktion	2/2	2/2
15. Beziehungen zwischen Zahlen	5/5	5/5
16. Approximativer Grössenvergleich	4/4	4/4
17. Subtraktion		10/10
	Total 81/83	Total 95/97

Beschreibung der Subtests und der Anpassungen gegenüber dem TEDI-MATH original

Um einen Deckeneffekt zu vermeiden, wurden nicht nur Subtests in den Test mitaufgenommen, die für den Kindergarten vorgesehen wurden, sondern auch solche für höhere Klassen. Um Kinder mit wenig Deutschkenntnissen nicht zu benachteiligen, wurde der Subtest *Entscheidung Zahlwort* aus der Originalversion nicht in den Test integriert. Hier werden den Kindern Wörter vorgesagt und sie müssen entscheiden, ob es sich um ein Zahlwort handelt oder nicht (z. B. zweizehn). Aus demselben Grund wurde auch auf den Subtest *Textaufgaben* verzichtet, bei dem den Kindern eine Aufgabe zur Addition bzw. zur Subtraktion vorgelesen wird. Zudem wurden bei Subtests mit sehr vielen Items einige weggelassen, um die Testzeit gering zu halten. Im Folgenden werden die Subtests und ihre Anpassungen beschrieben.

Subtest 1: *Zahlwortreihe*: Die Kinder sollen von 1 so weit wie möglich zählen, von einer bestimmten Zahl aus weiter zählen, rückwärts zählen, in Schritten zählen und um eine Anzahl Schritte weiterzählen (nur bei t2 und t3). Zwei Items aus der Originalversion wurden weggelassen (Zählen mit einer Obergrenze, Zählen mit Unter- und Obergrenze).

Subtest 2: *Abzählen*: Den Kindern wird eine abgebildete Anzahl Tiere gezeigt, die sie zählen sollen. Im TEDI-Original werden hier jeweils leicht unterschiedliche Aufgaben zur Nachfrage gestellt (z. B. „Wie viele Hasen sind es, wenn du bei diesem mit Zählen anfängst?“). In der vorliegenden Studie wurden bei allen Zählobjekten die drei gleichen Fragen, analog der Löwenaufgabe im TEDI-MATH, gestellt: „Zähle bitte laut alle Hasen. Wie viele Hasen sind es? Du hast gerade 9 Hasen gezählt“. (Hasenbilder zudecken) „Wie viele Hasen habe ich versteckt?“ Anstelle der 13 Items wurden nur 10 Items einbezogen. Eine weitere Aufgabe in diesem Subtest ist die Herstellung einer gleichmächtigen Menge. Die Kinder werden aufgefordert, gleich viele Flaschendeckel auf eine Unterlage zu legen, wie es abgebildete Punkte hat.

Subtest 3: *Entscheidung arabische Zahl*: Hier müssen Zahlen und von anderen Zeichen (Buchstaben, Dollarzeichen) unterschieden werden.

Subtest 4: *Grössenvergleich arabische Zahlen*: Den Kindern werden zwei Zahlen vorgelegt und es wird nach der grösseren Zahl gefragt. Um die Ratewahrscheinlichkeit zu minimieren, wurde in der vorliegenden Untersuchung jedes Item zweimal, zeitlich verschoben, mit vertauschten Zahlen vorgelegt und nur als richtig bewertet, wenn beide Antworten korrekt waren. Es wurden nur acht der insgesamt zwölf Items der Originalversion verwendet.

Subtest 5: *Transkodieren* (Zahlen lesen): Dieser Untertest ist nicht für den Kindergarten vorgesehen, wurde aber in dieser Untersuchung dennoch miteinbezogen. Von den 18 Items der Originalversion wurden 15 aufgenommen.

Subtest 6: *Anzahlen ordnen nach numerischer Grösse*. Die Kinder werden aufgefordert, fünf Karten mit Bildern zwischen einem und neun Bäumen bezüglich der Anzahl aufsteigend in eine Reihenfolge zu legen. Danach müssen sie eine weitere Karte am richtigen Ort in die Reihe einfügen.

Subtest 7: *Zahlen ordnen nach numerischer Grösse*. Analog zu Subtest 6 sollen die Kinder Karten mit Zahlen aufsteigend ordnen. Dieser Untertest ist in der Originalversion nicht für den Kindergarten vorgesehen. Da der Subtest nur ein Item enthielt, wurde ein neues Item entwickelt. Wie bei Subtest 6 musste eine Zahlenkarte nachträglich in die Reihe eingefügt werden.

Subtest 8: *Zahlen-Grössen-Zuordnung*: Diese Aufgabe wurde neu entwickelt. Die Kinder mussten eine Karte mit einer Anzahl Punkte (dargestellt im Zehner- bzw. Zwanzigerfeld) der richtigen Zahl zuordnen. Gemäss dem Entwicklungsmodell von Krajewski (2008) stellt die Zuordnung von Zahlen zu Mengen einen Meilenstein der mathematischen Entwicklung dar (vgl. Abschnitt 4.2.1), die in der Originalversion des TEDI-MATH nicht überprüft wird.

Subtest 9: *Numerische Inklusion*: Dieser Untertest ist nur in der Gesamtbatterie für das letzte Halbjahr des zweiten Kindergartenjahres vorgesehen, wurde aber in dieser Untersuchung schon zum ersten Testzeitpunkt eingesetzt. Die Kinder wurden aufgefordert, sechs Plättchen in einen

Umschlag zu stecken. Danach mussten sie Fragen beantworten, wie „Du hast sechs Plättchen im Umschlag. Kannst du mir acht Plättchen aus dem Umschlag geben?". Die Antworten mussten jeweils begründet werden. Die Aufgabe wurde nur dann als richtig bewertet, wenn die Begründung zur Antwort passte.

Subtest 10: *Additive Zerlegung*: In der Originalversion wird den Kindern ein Bild von zwei mit Zahlen beschrifteten Weiden gezeigt. In der hier verwendeten Version wurden anstelle der Zahlen Schafe in die Weiden gezeichnet. Die Kinder sollten weitere Möglichkeiten der Verteilung der Anzahl Schafe auf zwei Weiden finden.

Subtest 11: *Rechnen mit Objektabbildungen*: Die Subtraktionsaufgaben dieser Aufgabe werden im TEDI-MATH mit Bildern unterstützt, in denen nur die Gesamtmenge sichtbar ist, zum Teil in unterschiedlichen Farben markiert (z. B. fünf verschiedenfarbige Bälle). Dies kann zu Verwirrungen führen kann, da Minuend und Subtrahend nicht korrekt abgebildet sind. Die Bilder wurden deshalb durch Abbildungen ersetzt, in denen Minuend und Subtrahend eindeutig sichtbar sind (z. B. fünf Ballone, zwei Ballone fliegen weg).

Subtest 12: *Addition*: Hier werden Aufgaben auf der formal-symbolischen Ebene gezeigt (z. B. 6 + 3) und vorgelesen „sechs und drei, wieviel sind das zusammen?". Von den 18 Items der Originalversion wurden 9 aufgenommen.

Subtest 13: *Unvollständige Addition:* Dieser Untertest ist nur in der Gesamtbatterie für das letzte Halbjahr des zweiten Kindergartenjahres vorgesehen, wurde aber in dieser Untersuchung mit einer Anpassung der Repräsentationsebene schon zum ersten Testzeitpunkt eingesetzt. Die Aufgaben werden im TEDI-MATH original auf der formal-symbolischen Ebene dargestellt (z. B. 4 + ... = 8). In vorliegender Untersuchung wurde diese Aufgabe mit Plättchen veranschaulicht. Beispielsweise: „Hier siehst du vier Plättchen. Wie viele muss ich dazu legen, damit es sechs sind?".

Subtest 14: *Unvollständige Subtraktion*: Dieser Subtest ist in der Originalversion des TEDI-MATH ab der ersten Klasse vorgesehen und wird analog zur unvollständigen Addition auf der formal-symbolischen Ebene dargestellt (z. B. 9 - ... = 1). Auch diese Aufgabe wurde mit Plättchen veranschaulicht. Beispiel: „Hier siehst du drei Plättchen. Wie viele muss ich wegnehmen, damit es eins sind?".

Subtest 15: *Beziehung zwischen Zahlen*: Diese Aufgabe wurde neu in den Test integriert. Gemäss dem Entwicklungsmodell von Krajewksi (2008) stellt das Verständnis von Relationalzahlen einen zentralen Aspekt des numerischen Verständnisses dar (vgl. Abschnitt 4.2.1). Da dieses in der Originalversion des TEDI-MATH nicht überprüft wird, wurden fünf Aufgaben wie „Hier siehst du die Zahl fünf. Welche Zahl ist um zwei grösser als fünf?“ aufgenommen.

Subtest 17: *Subtraktion*: Dieser Untertest ist nicht für den Kindergarten vorgesehen, wurde aber beim zweiten und dritten Testzeitpunkt in dieser Untersuchung einbezogen. Die Aufgaben werden auf der formal-symbolischen Ebene gezeigt (7 - 4) und vorgelesen: „sieben weg vier, wieviel ist das?“ Von den 15 Items der Originalversion wurden 10 aufgenommen.

Die Überprüfung dieser adaptierten Form des TEDI-MATH-Tests zur Eignung der längsschnittlichen Erfassung ist Teil dieser Untersuchung (siehe erste Forschungsfrage, Abschnitt 7.1.1). Die Ergebnisse dieser Prüfung werden in Abschnitt 8.1 besprochen. Deshalb werden hier keine Angaben zu den Gütekriterien gemacht.

7.3.2 Kognitive Fähigkeiten

Für die vorliegende Untersuchung wurde ein Test zur Erfassung der kognitiven Grundfähigkeiten benötigt, der möglichst frei von soziokulturellen und sprachlichen Einflüssen war und mit einer ganzen Kindergruppe durchgeführt werden konnte. Der CFT 1-R (Weiss & Osterland, 2012) erfasst die Grundintelligenz mit ausschliesslich sprachfreien Aufgaben für Kinder ab 5.3 Jahren. Eine Kurzform dieses Tests mit den Subtests Ähnlichkeiten und Matrizen wurde zum zweiten Messzeitpunkt eingesetzt. Der Test zur Erfassung der kognitiven Fähigkeiten wies über alle 30 Items eine gute Reliabilität von Cronbachs Alpha .82 auf.

7.3.3 Erfassung von Hintergrundvariablen

Zum ersten Messzeitpunkt hatten die pädagogischen Fachkräfte Angaben zu sich und zu den Kindern gemacht. Damit wurden die Anzahl der Dienstjahre und die Ausbildung der pädagogischen Fachkräfte, das Geschlecht und das Geburtsdatum der Kinder erfasst. Die Eltern der Kinder hatten mit ihrer Einverständniserklärung ebenfalls Angaben zu ihrem Kind getätigt. Erfasst wurden das Geburtsland des Kindes, die Geburtsländer der Eltern sowie die zu Hause gesprochenen Sprachen. Die hauptsächlich gesprochene Sprache zu Hause wurde als Erstsprache bezeichnet. Die Variable Migrationshintergrund wurde vergeben, wenn beide Eltern nicht in der Schweiz bzw. nicht in Deutschland und auch nicht in Österreich geboren wurden.

7.4 Stichprobe

Zur Gewinnung der Stichprobe in der Schweiz wurden Schulleitungen bzw. kantonale Verantwortliche in acht grösseren deutschsprachigen Kantonen angeschrieben, mit der Bitte, ihre Kindergärtnerinnen und Kindergärtner mit den Projektinformationen zu beliefern. Dazu standen Flyer online und in Papierform zur Verfügung und es gab zusätzlich eine Webseite mit Informationen zum Projekt. Bei den Kindergärten in der Schweiz handelte es sich um ausschliesslich staatlich geführte Kindergärten. In Deutschland wurden öffentlich finanzierte Kindertagesstätten aus dem Grossraum Kiel und Vechta angefragt. Auch hier standen Flyer und die Website zur Verfügung. Die pädagogischen Fachkräfte meldeten sich freiwillig für das Projekt. Entsprechend handelt es sich um eine Selbstselektionsstichprobe (Döring & Bortz, 2016). Eine systematische Stichprobenziehung mit Anspruch auf regionale Repräsentativität ist für die Beantwortung der Fragestellungen in dieser Studie nicht notwendig.

7.4.1 Stichprobenbeschreibung

Vor Projektbeginn wurde eine Poweranalyse mit dem Programm G*Power durchgeführt, um die notwendige Stichprobengrösse zu bestimmen (Faul, Erdfelder, Buchner & Lang, 2009). Für die Gruppenvergleiche im WILMA-Projekt (Kontrollgruppe und zwei Interventionsgruppen) wurde von einer Effektgrösse von $d = 0.30$ ausgegangen. Bei einer Teststärke von $1\text{-}\beta = .95$ und einem Signifikanzniveau von $\alpha = .05$ ergab sich eine benötigte Anzahl von $n = 30$ pädagogischen Fachkräften für jede der drei Gruppen. Damit sind auch mehrebenenanalytische Analysen möglich. Da von einem Schwund von etwa 10 % ausgegangen werden muss, wurde eine Rekrutierung von 100 pädagogischen Fachkräften, davon 50 aus der Schweiz und 50 aus Deutschland angestrebt. Erfreulicherweise konnten aber bedeutend mehr Klassen für die Teilnahme am Projekt gewonnen werden. Insgesamt nahmen 158 Klassen (69 Kindergärten aus der Schweiz und 89 Kindertagesstätten aus Deutschland) und 1145 Kinder (569 aus der

Schweiz und 549 aus Deutschland) an der Untersuchung teil. Dabei wurden nur Einrichtungen berücksichtigt, bei der eine überwiegende Betreuung von Kindern durch eine pädagogische Fachkraft über zwei Jahre möglich war. Diese Anfangsstichprobe verringerte sich zum dritten Testzeitpunkt um einige Klassen und Kinder. Insgesamt konnten 894 Kinder zu allen drei Zeitpunkten getestet werden. Ihre Daten wurden in die Analysen miteinbezogen. Tabelle 4 gibt einen Überblick über die Untersuchungsstrichprobe.

Tabelle 5: Beschreibung der Untersuchungsstichprobe mit vorhandenen Daten zu allen drei Testzeitpunkten

	Schweiz	Deutschland	Gesamt
Anzahl Klassen	67	74	141
Anzahl Kinder	523	371	894
Mädchen/Jungen	258/265	179/192	437/457
Alter J (SD)	5.25 (0.37)	5.12 (0.36)	5.20 (0.37)
Kognitive Fähigkeiten Pt. (SD)	10.27 (4.66)	8.93 (4.77)	9.75 (4.75)
Erstsprache in % Deutsch/nicht Deutsch	76.1/23.9	81.7/18.3	78.3/21.7

Die Stichprobe umfasste in der Schweiz 67 Klassen und 596 Kinder, in Deutschland waren es 74 Klassen und 549 Kinder. Beim ersten Testzeitpunkt waren die Kinder in beiden Ländern durchschnittlich ca. fünf Jahre und zweieinhalb Monate alt, wobei die Kinder in der Schweiz insgesamt älter waren ($F[1,889] = 24.62, p = .000, n = 891$). Die kognitiven Fähigkeiten waren in Deutschland geringer ($M = 8.93$) als in der Schweiz ($M = 10.27$) ($F[1,846] = 16.41$, $p = .000$, $n = 848$). Obwohl in der Schweiz fast ein Viertel der Kinder eine andere Erstsprache als Deutsch hatte und in Deutschland weniger als ein Fünftel, wurde dieser Unterschied statistisch nicht signifikant ($\chi^2_{(df=1)} = 3.49, p = .062, n = 867$). Erwartungsgemäss befanden sich in der Gesamtstichprobe mehr Kinder mit Erstsprache Deutsch als mit einer anderen Erstsprache ($\chi^2_{(df=1)} = 278.06, p = .000, n = 867$). Die Verteilung der Geschlechter unterschied sich sowohl in der Gesamtgruppe ($\chi^2_{(df=1)} = .447, p = .504, n = 894$) als auch zwischen den beiden Ländern nicht ($\chi^2_{(df=1)} = .063, p = .802, n = 894$).

7.4.2 Leistungsgruppen

Die mathematischen Leistungen der Kinder unterschieden sich zum ersten Testzeitpunkt sehr stark. Die Personenparameter (WLE) reichten von -5.45 bis 5.53 Logits. Um diesen Umstand zu berücksichtigen und die Entwicklung der mathematischen Kompetenzen von Kindern mit niedrigen, mittleren und hohen Kompetenzen zu vergleichen (Hypothesen zur vierten Forschungsfrage, vgl. Abschnitt 7.1.4), wurden drei Leistungsgruppen gebildet. Ausschlaggebend für die Gruppeneinteilung waren der Mittelwert (M = -1.02) und die Standardabweichung (SD = 1.45) der Personenparameter zum ersten Testzeitpunkt. Kinder mit einer Standardabweichung unter dem Mittelwert (WLE kleiner als -2.47 Logits) wurden in die Gruppe mit niedrigen Leistungen eingeteilt, und Kinder mit einer Standardabweichung über dem Mittelwert (WLE grösser als 0.43 Logits) in die Gruppe mit hohen Leistungen. Die Gruppe mit mittleren Leistungen hatten WLE-Werte zwischen einer Standardabweichung unter und einer über dem Mittelwert (WLE-Werte zwischen -2.47 und 0.43 Logits). Tabelle 5 gibt einen Überblick über die drei Leistungsgruppen.

Tabelle 6: Beschreibung der Leistungsgruppen

	niedrig	mittel	hoch
Anzahl Kinder	124	637	133
Mädchen/Jungen	57/67	310/327	53/80
Alter J (SD)	5.09 (0.35)	5.17 (0.37)	5.40 (0.32)
Kognitive Fähigkeiten Pt. (SD)	6.52 (3.66)	9.50 (4.32)	13.97 (4.73)
Erstsprache in % Deutsch/nicht Deutsch	57.4/42.6	80.9/19.1	85.5/14.5

Das Kriterium der Einteilung mit jeweils einer Standardabweichung über bzw. unter dem Mittelpunkt führte aufgrund der Normalverteilung dazu, dass die Kinder mit mittleren Leistungen (632 Kinder) die grösste Gruppe bildeten. Die Gruppen der Kinder mit hohen bzw. niedrigen Leistungen waren mit 133 bzw. 124 Kindern etwa gleich gross. Der Anteil der Jungen war in der Gruppe mit hohen Leistungen am grössten. Die Gruppen unterschieden sich signifikant bezüglich der Verteilung der Geschlechter ($\chi^2_{(df=2)} = 6.30, p = .043, n = 894$). Die Kinder mit niedrigen Leistungen stellten durchschnittlich die jüngste Gruppe ($M = 5.09$, $SD = 0.35$) und die Kinder mit hohen Leistungen die älteste Gruppe ($M = 5.40, SD = 0.32$) dar. Die Gruppen unterschieden sich auch signifikant bezüglich des Alters ($F[3,888] = 3.72$, p = .000, n = 891). Erwartungsgemäss zeigten sich auch Unterschiede bezüglich kognitiver Fähigkeiten ($F[2,848] = 94.48, p = .000, n = 848$). Durchschnittlich hatten die Kinder mit hohen mathematischen Leistungen die höchsten kognitiven Fähigkeiten (*M* = 13.97, *SD* = 4.37), gefolgt von den Kindern mit mittleren mathematischen Leistungen (*M* = 9.50, *SD* = 4.32). Kinder mit niedrigen mathematischen Leistungen besassen auch die niedrigsten kognitiven Fähigkeiten (*M* = 6.52, *SD* = 3.66). Post-hoc-Tests mit Bonferroni-Korrektur zeigten, dass sich alle Gruppen bezüglich der kognitiven Fähigkeiten unterschieden (alle *p* = .000). Der Prozentsatz von Kindern mit einer anderen Erstsprache als Deutsch war in der Gruppe von Kindern mit niedrigen Leistungen mit 42.6 % am höchsten und in der Gruppe mit hohen Leistungen am niedrigsten (14.5 %). Die Gruppen unterschieden sich signifikant bezüglich der Verteilung der Erstsprache ($\chi^2_{(df=2)}$= 37.97, *p* = .000, *n* = 867).

7.5 Statistische Methoden

Zur Beantwortung der Forschungsfragen wurden unterschiedliche Auswertungsmethoden eingesetzt. Zur Überprüfung des Mathematiktestinstrumentes wurden Analysen aus der klassischen und der probabilistischen Testtheorie durchgeführt. Diese wurden in den Abschnittn 6.1 und 6.2 beschrieben und werden daher hier nicht mehr aufgeführt. Deskriptive Analysen erfolgten, um die mathematischen Kompetenzen der Kinder beim ersten Testzeitpunkt zu beschreiben. Um Zusammenhänge zwischen verschiedenen Variablen aufzuzeigen, wurden zudem Korrelationen berechnet. Zur Prüfung von Mittelwertunterschieden der verschiedenen Gruppen wurden Varianzanalysen eingesetzt. Ein besonderes Interesse dieser Studie galt auch den mittleren Veränderungswerten der Mathematikleistungen der Kinder über die drei Messzeitpunkte und die individuellen Unterschiede in diesen Mittelwertveränderungen. Zudem interessierten mögliche Einflussfaktoren zur Erklärung dieser Unterschiede. In den letzten Jahren haben sich Verfahren von Veränderungsmessungen etabliert, in denen feste und zufällige Effekt kombiniert werden. „Die festen Effekte repräsentieren die mittleren Veränderungen und die zufälligen Effekte individuelle Abweichungen von diesen. Zu diesen Verfahren zählen sowohl Mehrebenenmodelle (Multilevel models bzw. Mixed models) als auch latente Wachstumskurvenmodelle (Latent growth curve models; LGMs)“ (Schmiedek & Wolff, 2010, S. 1). Beide statistischen Analyseverfahren wurden zur Beantwortung der Fragestellungen in dieser Studie angewandt. Folgend werden diese Verfahren erläutert.

7.5.1 Deskriptive Analysen und Gruppenunterschiede

Es wurden Analysen zur Beschreibung der Stichproben und der mathematischen Kompetenzen der Kinder zu den verschiedenen Testzeitpunkten durchgeführt. Dabei wurden sowohl Mittelwerte und Standardabweichungen über den Gesamttest berichtet als auch prozentuale Angaben darüber, welche Aufgaben bzw. Aufgabengruppen zu verschiedenen Zeitpunkten gelöst wurden. Mittels Chi-Quadrat-Tests wurden Zusammenhänge von kategorialen Variablen wie Geschlecht und Erstsprache zwischen Gruppen (beispielsweise Kinder in der Schweiz und in Deutschland oder Kinder mit niedrigem, mittlerem und hohem Vorwissen) aufgezeigt und Aussagen über die Verteilung getroffen. Die Voraussetzungen der Stichprobengrösse (> 50) und der erwarteten Zellhäufigkeit > 5) waren erfüllt. Dort wo die Freiheitsgrade nicht grösser als 1 waren, wurde die Korrektur nach Yates verwendet (Bortz & Schuster, 2016). Unterschiede zwischen Gruppen und mit intervallskalierten Variablen (kognitive Fähigkeiten und Alter) wurden mittels einfaktorieller Varianzanalysen berechnet. Diese werden in Abschnitt 7.5.5 beschrieben.

7.5.2 Korrelationen

Für die Analyse verschiedener Zusammenhänge wurde die Produkt-Moment-Korrelation nach Pearson (r) berechnet. Die Korrelationskoeffizienten geben dabei an, wie stark der Zusammenhang zwischen zwei Merkmalen ist. Nach der von Cohen (1992) formulierten Konvention gelten Werte ab $r = .10$ als schwache Korrelation, $r = .30$ als mittlere und $r = .50$ als starke Korrelation.

7.5.3 Mehrebenenanalyse

Mit Mehrebenenanalysen können geschachtelte Daten ausgewertet werden. Die in dieser Studie vorliegenden Daten wiesen eine hierarchische Struktur auf, da die Kinder in Klassen geschachtelt bzw. geclustert waren. Unterschieden wird im Allgemeinen zwischen Level-1- und Level-2-Einheiten. Die einzelnen Kinder sind der Individualebene (Level 1) zuzuordnen. Die zweite Ebene (Level 2) umfasst die Klassen, denen die Kinder angehören. Die Berücksichtigung der hierarchischen Struktur von Daten ist insofern bedeutsam, als dadurch die Annahme der Unabhängigkeit der Beobachtungseinheiten nicht verletzt wird. Bezogen auf die vorliegende Stichprobe wäre es beispielsweise möglich, dass sich Kinder ein- und derselben

Klasse bezüglich der mathematischen Leistung ähnlicher sind als Kinder, die zu einer anderen Klasse gehören, beispielsweise weil die Fachkraft die mathematische Förderung im Kindergarten als sehr wichtig betrachtet. Die Beachtung der Clusterung in Gruppen verhindert auch eine Verzerrung der Ergebnisse durch eine Überschätzung der Stichprobengrösse. Zudem können verschiedene Prädiktoren auf der zu ihnen passenden Ebene adäquat eingesetzt werden (Geiser, 2011). Aus diesem Grund wurden zur Erklärung der Unterschiede in den Leistungen und der Kompetenzentwicklung der Kinder multiple, lineare Regressionen berechnet, die die Mehrebenenstruktur mitberücksichtigen. Die vorliegenden Daten erfüllten auch die Voraussetzungen, die generell für das Berechnen von Regressionen notwendig sind. Es sind dies eine intervallskalierte abhängige Variable, intervallskalierte oder dichotome unabhängige Variablen, ein linearer Zusammenhang zwischen abhängigen und unabhängigen Variablen, gleiche Varianzen der Residuen (Homoskedastizität) und Unkorreliertheit der Residuen. Zur Vermeidung von Multikollinearität sollten die unabhängigen Variablen nicht zu stark korrelieren (Bortz & Schuster, 2016). Aus diesem Grund wurde der Migrationshintergrund nicht in die Analysen miteinbezogen, weil er mit der Erstsprache hoch korreliert (vgl. Abschnitt 7.1.2).

In Zwei-Ebenen-Modellen wird zwischen dem Random-Intercept-Modell und dem Random-Slope-Modell differenziert. Beim Random-Intercept-Modell wird von unterschiedlichen Ausgangsleistungen (Intercepts) ausgegangen und überprüft, ob Varianzen auf der Ebene 1 und der Ebene 2 vorhanden sind. Beim Random-Slope-Modell hingegen interessiert, ob Varianzen in den Intercepts und Slopes vorhanden sind. Zur Überprüfung, welches Modell die Daten besser erklärt, erfolgt ein Modellvergleich mit ANOVA. Dabei werden im Random-Slope-Modell sowohl Intercept als auch Slope frei geschätzt, während im Random-Intercept-Modell nur der Intercept frei geschätzt wird und der Slope fixiert ist. Es werden neben der statistischen Signifikanz auch Modellgüte-Werte (AIC und BIC) berechnet. AIC und BIC sind informationstheoretische Masse, die zum Vergleich verschiedener Modelle dienen. Das Modell mit den kleinsten Werten passt jeweils besser zu den Daten (Geiser, 2011). In der vorliegenden Untersuchung wurden zu allen drei Testzeitpunkten jeweils beide Modelle berechnet und mittels einer Varianzanalyse verglichen. Diese und alle anderen Mehrebenenanalysen wurden mit Hilfe der Statistiksoftware R und dem Package *lavaan* (Rosseel, 2017) berechnet.

Oft werden bei der Analyse mit Mehrebenenmodellen Prädiktoren schrittweise ins Modell aufgenommen, um zu überprüfen, ob sich das Modell verbessert bzw. die erklärte Varianz vergrössert wird. In dieser Studie wurde der Einfluss von Prädiktoren nicht explorativ erforscht, sondern hypothesengeleitet. Das heisst, es wurden jeweils die Prädiktoren ins Modell aufgenommen, die aufgrund der Hypothesen auf ihre Bedeutung hin untersucht werden sollten. Um die Bedeutung der Koeffizienten besser interpretieren zu können, wurden alle nicht dichotomen Variablen am Gesamtmittelwert zentriert.

Zur Beurteilung von Mehrebenenmodellen wird zum einen die Aufklärung der Gesamtvarianz durch das Modell berichtet. In R wird die Gesamtvarianz mit dem R^2 aus dem Package *lme4* (Bolker, 2018) berechnet. Dabei stehen ein marginal und ein conditionel R^2_{GLMM} zur Verfügung, wobei der marginale Wert aussagekräftiger ist (Nakagawa & Schielzeth, 2013). Zusätzlich ist auch von Interesse, den erklärenden Varianzanteil eines Prädiktors zu bestimmen und so eine inhaltliche Aussagekraft der jeweiligen Variablen zu erhalten. Die klassische Berechnungsweise der Varianzaufklärung in Mehrebenenmodellen, wie sie beispielsweise bei Hox, Moerbeek und Schoot (2017) zu finden ist, berechnet sich aus der Differenz der Varianz mit und ohne den interessierenden Prädiktor auf der jeweiligen Ebene:

$$R^2 = \left(\frac{\sigma 2Nullmodell - \sigma 2Vergleichsmodell}{\sigma 2Nullmodell} \right)$$

In vorliegender Untersuchung wurde der erklärte Varianzanteil durch einzelne Prädiktoren mit dieser Formel berechnet.

7.5.4 Varianzanalyse

Die Mehrebenenanalysen zeigen unter anderem den Einfluss von Kontextfaktoren auf die Leistungsentwicklung der Kinder zwischen dem ersten und zweiten Testzeitpunkt bzw. zwischen dem zweiten und dritten Testzeitpunkt. Darüber hinaus können auch Aussagen darüber gemacht werden, ob bestimmte Merkmale wie Geschlecht, Erstsprache oder Landzugehörigkeit die Leistungsentwicklung beeinflussen. Sie geben aber keinen Aufschluss darüber, ob sich die verschiedenen Gruppen zum zweiten bzw. zum dritten Testzeitpunkt in ihrer Kompetenz unterscheiden. Deshalb wurden zusätzlich einfaktorielle Varianzanalysen (ANOVA) gerechnet. Die Anwendung ist an die Voraussetzung der Unabhängigkeit der Beobachtungen, der Normalverteilung, der Intervallskalierung sowie der Varianzhomogenität gebunden (Bortz & Schuster, 2016). Die Voraussetzung der Unabhängigkeit der in dieser Studie gebildeten Gruppen (Mädchen und Jungen, Kinder aus der Schweiz und aus Deutschland, Kinder mit Erstsprache Deutsch und Kinder mit einer anderen Erstsprache) war gegeben. Zudem war die abhängige Variable (mathematische Leistung) intervallskaliert. Die Normalverteilung innerhalb der Gruppen und die Homogenität der Varianzen waren teilweise leicht verletzt. Die Varianzanalyse gilt jedoch bei grossen Stichproben mit je mehr als zehn Probanden als robust gegenüber diesen Verletzungen (Bortz & Schuster, 2016).

7.5.5 Latente Wachstumskurvenmodelle (LGCM)

Die latenten Wachstumskurvenmodelle wurden in dieser Untersuchung eingesetzt, um den Entwicklungsverlauf der Gesamtstichprobe über die drei Messzeitpunkte zu analysieren und um unterschiedliche Entwicklungsverläufe und Einflussfaktoren von Kindern mit niedrigem, mittlerem und hohem Vorwissen zu vergleichen. Latente Wachstumskurvenmodelle (latent growth curve models, LGCM) zeigen die mittleren Veränderungen und interindividuellen Unterschiede mit einer Kombination von festen und zufälligen Effekten an, wobei die festen Effekte die mittleren Veränderungen und die zufälligen Effekte individuelle Abweichungen von diesen repräsentieren (Schmiedek & Wolff, 2010). Es handelt sich um Spezialfälle von Strukturgleichungsmodellen (SGM). Mit latenten Wachstumskurvenmodellen lassen sich der mittlere Ausgangswert (intercept) des interessierenden Merkmals und dessen Veränderung über die Zeit (mittlere Anstieg: slope) berechnen und zudem kann bestimmt werden, in welcher Form die Veränderung stattfindet (z. B. lineares oder quadratisches Wachstum). Des Weiteren können individuelle Unterschiede im Wachstum und der Zusammenhang von Ausgangs- (intercept) und Wachstumswert (slope) berechnet werden und es lässt sich so feststellen, ob gewisse Variablen (Kovariaten) Unterschiede im Ausgangs- bzw. im Wachstumswert erklären können (Geiser, 2011). Ein Vorteil gegenüber herkömmlichen Verfahren zur Analyse von Längsschnittdaten, wie beispielsweise der Varianzanalyse mit Messwiederholung (ANOVA), besteht darin, dass unterschiedliche Zeitabstände in den Messungen berücksichtigt werden können. Eine grundlegende Unterscheidung erfolgt dabei zwischen Wachstumskurvenmodellen erster und zweiter Ordnung. LGCM erster Ordnung enthalten einen einzelnen wiederholt gemessenen Indikator, während LGCM zweiter Ordnung pro Messzeitpunkt mehrere, gleiche Indikatoren haben. In der vorliegenden Untersuchung wurden LGCM erster Ordnung berechnet, da nur ein messwiederholender Indikator (mathematische Kompetenz) pro Messzeitpunkt vorkam.

In der Literatur werden verschiedene Kriterien zur Beurteilung der Modellgüte von Wachstumskurvenmodellen bzw. Strukturgleichungsmodellen genannt. Eine gute Passung bedeutet, dass die mittels des Modells rekonstruierte Datenstruktur von der empirisch gegebenen Datenstruktur nur geringfügig abweicht. Folgend werden die gängigen Tests zur Beurteilung von Modellpassungen und Cut-Werte beschrieben (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger & Müller, 2003).

Chi-Quadrat-($\mathcal{X}^2$) Test: Der Chi-Quadrat-Test testet die Nullhypothese, dass die Datenstruktur des Modells gleich der empirisch beobachteten ist. Da der $\mathcal{X}^2$-Wert sehr sensitiv gegenüber grossen Stichproben ist, können bereits kleinste Abweichungen zwischen modellbasierter und beobachteter Kovarianzmatrix zu einer Ablehnung der Nullhypothese führen. Es empfiehlt sich daher, bei grossen Stichproben ($n > 500$) auf andere Indizes zur Beurteilung der Modellgüte zurückzugreifen (Schermelleh-Engel et al., 2003). Deshalb wurden in dieser Untersuchung mit einer Stichprobengrösse von $n = 894$ andere Gütemasse berücksichtigt.

Root Mean Square Error of Approximation (RMSEA): Der RMSEA ist ein Mass für den approximativen Datenfit und orientiert sich am saturierten Modell. Der Wert sollte möglichst klein sein. Werte kleiner oder gleich .05 können als guter Fit betrachtet werden. Werte zwischen .05 und .08 sind akzeptabel, Werte zwischen .08 und .10 werden als mittelmässig bezeichnet. Nicht akzeptabel sind Werte über .10 (Browne & Cudeck, 1989).

Standardized Root-Mean-Residual (SRMR): In Ergänzung zum RMSEA gibt der SRMR die standardisierte durchschnittliche Abweichung zwischen beobachteter und modellbasierter Korrelationsmatrix an. Auch dieser Wert sollte möglichst klein sein. Werte kleiner .05 gelten als gut, während Werte grösser .10 als akzeptabel interpretiert werden können (Hu & Bentler, 1995).

Comparative-Fit-Index (CFI): Der CFI ist ein relativer Fit-Index, der sich nicht am saturierten Modell, sondern am Nullmodell orientiert. Er liegt zwischen 0 und 1 und sollte möglichst nahe bei 1 liegen. Werte grösser oder gleich .97 gelten als gut und Werte grösser oder gleich .95 sind akzeptabel (Schermelleh-Engel et al., 2003).

In der vorliegenden Untersuchung wurden jeweils der RMSEA-, SRMR- und der CFI-Wert berechnet und zur Prüfung der Modellgüte beigezogen.

8 Ergebnisse

In diesem Kapitel werden nun die Ergebnisse der Analysen wiedergegeben, die zur Beantwortung der Fragestellungen hinleiten und eine Diskussion der Hypothesen ermöglichen. In Abschnitt 8.1 werden zuerst die Ergebnisse der Prüfung des Mathematiktests mit dem Rasch-Modell aufgezeigt. Danach werden in Abschnitt 8.2 die Ergebnisse der Analysen der mathematischen Kompetenzen zum ersten Testzeitpunkt berichtet. Im Zentrum stehen dabei die Fragen nach der Verteilung der mathematischen Kompetenzen und nach den Einflussfaktoren. Abschnitt 8.3 widmet sich dann den Ergebnissen der Analysen zur mathematischen Leistungsentwicklung. Zudem werden die Ergebnisse bezüglich der Leistungsentwicklung der Gesamtstichprobe und zu den Einflussfaktoren auf diese Entwicklung dargelegt. Anschliessend werden in Abschnitt 8.4 die Entwicklungsverläufe von unterschiedlichen Gruppen (Mädchen und Jungen, Kinder mit Erstsprache Deutsch und mit einer anderen Erstsprache, Kinder in der Schweiz und in Deutschland) beschrieben und graphisch dargestellt. Abschnitt 8.5 fokussiert danach die Ergebnisse der Analysen zu den Unterschieden im Entwicklungsverlauf von Kindern mit niedrigem, mittlerem und hohem Vorwissen. Eine Zusammenfassung der wichtigsten Ergebnisse im Überblick schliesst dieses Kapitel ab.

8.1 Überprüfung des Mathematiktests mit dem Rasch-Modell

Da nur zwei Items von insgesamt 81 bzw. 95 Items eine zweistufige Antwortkategorie (partial credit) hatten, wurden diese umcodiert und mit richtig oder falsch gewertet. Dadurch konnten die Analysen zur Überprüfung des Mathematiktests mit dem (dichotomen) Rasch-Modell gerechnet werden (vgl. Abschnitt 6.2.1). Im Folgenden werden die Ergebnisse dieser Überprüfung mit dem Rasch-Modell präsentiert. Dabei werden zuerst die Ergebnisse zu den Item-Fitwerten vorgestellt, gefolgt von den Analyseresultaten zur Mess- und Subgruppeninvarianz. Eine Tabelle zeigt die ausgeschlossenen Items im Überblick. Schliesslich werden die Ergebnisse zur Schätzung der Personenparameter mittels mehrdimensionaler Raschskalierung dargestellt.

8.1.1 Item-Fit-Werte

Zur Itemanalyse wurde der Item-Fit (Infit-MNSQ, vgl. Abschnitt 6.2.3) mit Hilfe der Statistiksoftware R (Version 3.5.3) und dem Package ‚TAM' von Robitzsch, Kiefer und Wu (2017) bestimmt. Dabei wurden eindimensionale Raschmodelle für alle drei Testzeitpunkten berechnet. Insgesamt wurden 81 Ankeritems über drei Testzeitpunkte und zusätzlich 14 Items über den zweiten und dritten Testzeitpunkt in die Analysen einbezogen. Für die Überprüfung der Item-Fit-Werte wurden in dieser Untersuchung, analog zu Wilson (2005), Items mit einem Infit-MNSQ-Wert zwischen 0.75 und 1.33 toleriert. Dies bedeutet, dass eine maximal 25 % geringere bzw. eine maximal 33 % höhere Variation zwischen den beobachteten und den erwarteten Antwortmustern gegenüber einer perfekten Modellpassung vorliegt. Beim ersten Testzeitpunkt wiesen 76 von 81 Items einen MNSQ-Wert im tolerierten Bereich auf. Diese Items entsprachen also den Anforderungen. Zwei Items hatten einen leicht zu hohen und zwei Items einen leicht zu niedrigen Fit. Beim zweiten Testzeitpunkt wiesen 92 von insgesamt 95 Items einen Wert zwischen 0.75 und 1.33 auf. Auch hier hatten wieder zwei Items einen leicht zu hohen Fit und ein Item hatte einen leicht zu niedrigen Wert. Beim dritten Testzeitpunkt wiesen drei Items einen leicht zu hohen und zwei Items einen zu niedrigen Fit auf. Da leicht zu niedrige MNSQ-Werte, sogenannte Overfits, im Allgemeinen als unproblematisch beschrieben werden (Rost, 2004), wurden sie nicht aus den Testanalysen ausgeschlossen. Es wurden nur die vier Items mit einem Infit-MNSQ über 1.33 ausgeschlossen (vgl. Tabelle 8).

8.1.2 Überprüfung der Messinvarianz

Zur Überprüfung der Messinvarianz wurden ebenfalls mit Hilfe der Statistiksoftware R und dem Package ‚TAM' von Robitzsch, Kiefer und Wu (2017) DIF-Analysen (Differential Item Functioning) zur zeitbezogenen Messinvarianz und zur Subgruppeninvarianz durchgeführt (vgl. Abschnitt 6.2.3).

Messinvarianz über die Zeit

Zuerst erfolgte eine DIF-Analyse auf Itemebene unter Einbezug aller drei Testzeitpunkte. Diese Überprüfung zeigte, dass die Abweichungen der einzelnen Items von der mittleren Itemschwierigkeit über alle drei Testzeitpunkte sehr gering waren. Die Korrelationen nach Pearson betrugen zwischen dem ersten und zweiten Testzeitpunkt .98, zwischen dem zweiten und dritten ebenfalls .98 und zwischen dem ersten und dritten Testzeitpunkt .96. Dies bedeutet, dass der Anteil geteilter Varianz über die Testzeitpunkte hinweg 92 % beträgt und davon ausgegangen werden kann, dass die Rangreihenfolge der Itemschwierigkeit zu den einzelnen Testzeitpunkten sehr stabil war. Trotzdem wurden noch zusätzlich DIF-Analysen bezüglich der Testzeitpunkte auf Itemebene berechnet. Im Ergebnis wurde dabei sichtbar, dass insgesamt 20 Items zwischen zwei Testzeitpunkten eine mässige bis hohe Form von DIF aufwiesen (vgl. Tabelle 7). Die Unterschiede waren höher als 0.639 Logits (vgl. Abschnitt 6.2.3). Erwartungsgemäss waren Itemvergleiche zwischen dem ersten und dritten Testzeitpunkt stärker von DIF betroffen. Die betroffenen Items wurden aus den Analysen ausgeschlossen.

Subgruppeninvarianz

Zur Überprüfung der Personenhomogenität fanden ebenfalls DIF-Analysen statt. Als Teilungskriterium wurden das Geschlecht und das Land herangezogen. So konnte überprüft werden, ob für Mädchen und Jungen bzw. für Kinder in der Schweiz und Kinder in Deutschland mit jeweils ähnlichen Personenfähigkeitsparametern dieselbe Wahrscheinlichkeit bestand, ein bestimmtes Item zu lösen. Die DIF-Analysen zum Land ergaben, dass einige Items in den beiden Ländern unterschiedliche Itemschwierigkeiten aufwiesen. Beim ersten Testzeitpunkt waren es 3 Items, beim zweiten 7 und beim dritten 13. Die Unterschiede zwischen den Ländern dieser Items lagen zwischen 0.69 und 1.20 Logits, was eine mässige bis hohe Form von DIF bedeutet (vgl. Abschnitt 6.2.3). Bei den DIF-Analysen zu Unterschieden bei Jungen und Mädchen zeigte sich ein ähnliches Bild. Auch hier waren einige Items von mässiger bis hoher Form von DIF betroffen (Unterschiede zwischen den Geschlechtern zwischen 0.70 und 2.60 Logits). Beim ersten Testzeitpunkt waren es 6, beim zweiten 7 und beim dritten Testzeitpunkt 12. Da in dieser Studie auch die unterschiedliche mathematische Kompetenzentwicklung in den beiden Ländern und von Mädchen und Jungen untersucht werden sollte, wurden sämtliche Items, die zu irgendeinem Zeitpunkt eine mässige bis hohe Form von DIF aufwiesen, aus den Analysen ausgeschlossen.

Tabelle 7 gibt einen Überblick über die Ergebnisse der Überprüfung von Messinvarianz über die drei Testzeitpunkte und bezüglich Subgruppen.

Tabelle 7: Ergebnisse der Überprüfung von Messvarianz, DIF in Logits nach Paek und Wilson (2011)

Logits	$p \leq 0.05$	Mess(in)varianz: Anzahl Items		
		Zeit	Land	Geschlecht
DIF ≥ 0.638	mässig bis hoch	19	22	14
0.426 < DIF < 0.638	leicht bis mässig	24	27	21
DIF ≤ 0.426	vernachlässigbar	52	46	60

8.1.3 Überblick über alle ausgeschlossenen Items

Insgesamt wurden aufgrund von schlechtem Item-Fit oder Messinvarianz beim ersten Testzeitpunkt 35 und beim zweiten und dritten Testzeitpunkt 41 Items aus den Analysen ausgeschlossen. Tabelle 8 gibt einen Überblick über die ausgeschlossenen Items.

Tabelle 8: Übersicht ausgeschlossener Items

Subgruppe	Item	Schlechter Item-Fit	Messvarianz über die Zeit	Messvarianz Land	Messvarianz Geschlecht
Zahlwortreihe	Vorwärts 3 bis 10 zählen			x	
	Rückwärts von 15		x		x
	Zweierschritte vorwärts		x		
	Zehnerschritte rückwärts		x		
	Von 8 um 5 Schritte weiter			x	
	Von 9 um 6 Schritte weiter			x	
Abzählen	5 Objekte abzählen				x
	Nachfrage 1: Wie viele?		x	x	
	Nachfrage 2 nach Abdecken: Wie viele?		x		x
	9 Objekte abzählen		x	x	
	Nachfrage 1: Wie viele?			x	
	Nachfrage 2 nach Abdecken: Wie viele?		x		
	Nachfrage 2 bei 12 Objekten nach Abdecken: Wie viele?			x	
Grössenvergleich	109/180	x	x		
	403/420	x	x		
Zahlen lesen	3		x		x
	30		x	x	x
	15				x
	47		x		
	80				x
	92			x	x
	105				x
	800			x	x
	160				x
	Karten mit einer Anzahl Bäume nach Grösse ordnen			x	

Anzahlen nach numerischer Grösse ordnen	Eine Baumkarte richtig einordnen			x	
Zahlen nach numerischer Grösse ordnen	Karten mit Zahlen nach Grösse ordnen			x	
	Eine Zahlenkarte richtig einordnen			x	
Zahlen-Grössen-Zuordnung	Zahlen-Grössen-Zuordnung				x
Numerische Inklusion	Im Briefumschlag sind 6 Plättchen. Kannst du mir 4 daraus geben?	x	x		
Addition	2 + 2		x		
	20 + 8		x		
	32 + 14			x	x
	9 - 5		x	x	
Subtraktion (17)	7 - 4			x	
	4 - 0			x	
	27 - 6			x	x
Approx. Punktevergleich (16)	Punktevergleich 4 und 6		x		
	Punktevergleich 7 und 2	x	x		
	Punktevergleich 7 und 12		x		
	Punktevergleich 15 und 8		x		

Schlechte Fit-Werte wiesen Items mit einer hohen Itemschwierigkeit (Grössenvergleich) bzw. Items mit einer 50-prozentigen Ratewahrscheinlichkeit (Numerische Inklusion und Punktevergleich) auf. Die Items, die wegen Messvarianz über die Zeit ausgeschlossen wurden, waren insbesondere Items zum Zählen und Abzählen, zum Grössenvergleich, zum Zahlen lesen, zur Addition und zum approximativen Punktevergleich. Von länderspezifischer Messvarianz waren alle Items zum Anzahlen bzw. Zahlen nach numerischer Grösse ordnen betroffen, aber auch Items zum Zählen, zum Abzählen und zur Addition und zur Subtraktion. Geschlechterspezifische Messvarianz zeigten vor allem Items zum Transkodieren von Zahlen. Mögliche Gründe für schlechte Fit-Werte bzw. für Messvarianz werden in Abschnitt 9.2 diskutiert.

Tabelle 9 stellt diejenigen Items dar, die für die verschiedenen Analysen zur Beantwortung der Fragestellungen verwendet wurden und auch die Grundlage für die Schätzung der Personenparameter bildeten.

Tabelle 9: Übersicht über Subtests, mathematische Aktivität und Anzahl Items

Subtest	Mathematische Aktivität	Anzahl Items T1	Anzahl Items T1/T2
Zahlwortreihe	Zählen vorwärts und rückwärts	3	3
Abzählen	Anzahl abgebildeter Gegenstände bestimmen, zwei numerisch äquivalente Mengen herstellen	3	3
Entscheidung arabische Zahl	Zahlen von anderen Zeichen unterscheiden	1	1
Grössenvergleich arabische Zahl	Zahlen vergleichen: Welche Zahl ist grösser?	6	6
Transkodieren	Zahlen lesen	6	6
Numerische Inklusion	Anzahlen vergleichen mit nicht sichtbarer Menge	2	2
Additive Zerlegung	Zahlen zerlegen und ergänzen	2	4
Rechnen mit Objektabbildungen	Mengen zusammensetzen und verringern	6	6
Addition	Addieren auf formaler Ebene	6	6
Unvollständige Addition	Ergänzen, gestützt mit Plättchen	4	4
Unvollständige Subtraktion	Verringern, gestützt mit Plättchen	2	2
Beziehungen zwischen Zahlen	Relationalzahl	5	5
Subtraktion	Subtrahieren auf formaler Ebene	--	6
Total		46	54

8.1.4 Schätzung der Personenparameter

Um Aussagen über den Lernzuwachs der Kinder machen zu können, wurden Veränderungsschätzungen auf Basis der verbliebenen 46 bzw. 54 Items ebenfalls mit der Statistiksoftware R und dem Package ‚TAM' durchgeführt. Als Personenschätzer wurden die gängigen WLE-Personenparameter (Warms Weighted Likelihood Estimate) mittels mehrdimensionaler Raschskalierung geschätzt (vgl. Abschnitt 6.2.2). Die WLE-Reliabilität war zu allen Zeitpunkten hoch (vgl. Tabelle 10). Die Werte der Personenfähigkeit ergeben sich aus der auf einen Mittelpunkt von 0 fixierten Itemschwierigkeit und einer 50-prozentigen Lösungswahrscheinlichkeit, wobei die Itemschwierigkeit der identischen Items über die Testzeitpunkte gleichgesetzt wurde. Sie lagen beim ersten Testzeitpunkt zwischen -5.45 und

5.53 mit einem Mittelwert von $M = -1.02$ ($SD = 1.45$), beim zweiten Messzeitpunkt zwischen - 5.45 und 8.39 mit einem Mittelwert von $M = -.22$ ($SD = 1.64$) und beim dritten Testzeitpunkt zwischen -3.02 und 8.37. Der Mittelwert betrug hier $M = 1.27$ und die Standardabweichung $SD = 1.69$. Die Korrelation der Personenfähigkeitsparameter zwischen der Testversion mit allen Items und der Version mit den reduzierten Items betrug zu allen drei Testzeitpunkten $r = .98$. Es kann somit mit hoher Wahrscheinlichkeit angenommen werden, dass beide Versionen das gleiche Konstrukt gemessen haben.

Tabelle 10 fasst die Ergebnisse der Reliabilitätsprüfung und des Item-Fits nach Ausschluss der problematischen Items zusammen.

Tabelle 10: Übersicht Cronbachs Alpha, MNSQ-Werte und WLE-Reliabilität

	T1	T2	T3
Cronbachs Alpha	.93	.94	.94
Infit-MNSQ	0.75–1.27	0.75–1.31	0.74–1.21
WLE-Reliabilität	.92	.94	.93

Cronbachs Alpha war zu allen drei Testzeitpunkten hoch (zwischen .93 und .94). Der Item-Fit lag bei allen drei Testzeitpunkten innerhalb der tolerierten Grenzen zwischen 0.75 und 1.31. Auch die Personenreliabilität war hoch (zwischen .92 und .94).

Boden- oder Deckeneffekt

Da die mathematischen Kompetenzen zu drei Testzeitpunkten gemessen wurden, war es wichtig, einen Boden- bzw. Deckeneffekt zu vermeiden (vgl. Abschnitt 7.3.1). Ein Bodeneffekt (auch floor effect genannt) ist dann vorhanden, wenn Personen mit geringen Testleistungen keine Aufgaben lösen können. Entsprechend besteht ein Deckeneffekt (auch ceiling effect genannt), wenn der Test für Personen mit hohen Fähigkeiten zu einfach ist und viele Personen alle Items lösen können (Döring & Bortz, 2016). Die Testauswertungen des hier eingesetzten Mathematiktests haben gezeigt, dass kein Boden- und Deckeneffekt vorhanden war. Es gab bei keinem Testzeitpunkt Kinder ohne richtige Lösung und lediglich drei Kinder von 894 erreichten beim dritten Testzeitpunkt die maximale Punktzahl.

8.2 Mathematische Kompetenzen beim ersten Testzeitpunkt und deren Einflussfaktoren

Im Folgenden werden die Hypothesen zur zweiten Forschungsfrage (siehe Abschnitt 7.1.2) überprüft. Im Zentrum der Auswertungen stand die Frage, über welche mathematischen Kompetenzen die Kinder beim ersten Testzeitpunkt verfügen und welche Einflussfaktoren für die Leistungsunterschiede bedeutsam waren.

8.2.1 Mathematische Kompetenzen beim ersten Testzeitpunkt

Die Leistungen der Kinder wurden mit dem TEDI-MATH (vgl. Abschnitt 7.3.1) erhoben. Der Mittelwert der Personenparameter (WLE) der Untersuchungsstichprobe (N = 894) betrug M = -1.02 mit einer Standardabweichung von SD = 1.45, was auf eine grosse Streuung hinweist. Die Werte der Personenparameter variierten zwischen -5.45 Logits (niedrigster Personenparameter) und 5.53 Logits (höchster Personenparameter).

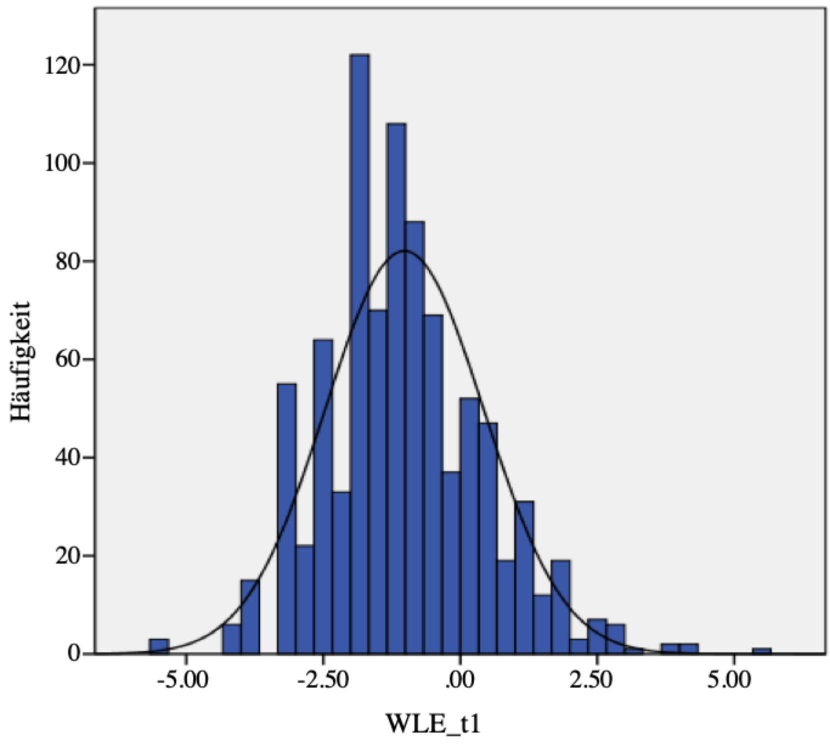

Abbildung 4: Mathematische Kompetenzen zum ersten Testzeitpunkt

Das einfachste Item lösten 80.5 % der Kinder beim ersten Testzeitpunkt. Es handelte sich um eine Aufgabe zur unvollständigen Subtraktion mit kleinen Zahlen und mit Veranschaulichung: „Hier siehst du 3 Plättchen. Wie viele muss ich wegnehmen, damit es 1 sind?“ Betrachtet man die zehn einfachsten Items, so lösten diese durchschnittlich 63.5 % aller Kinder. Es handelte sich neben dem eben beschriebenen Item zur unvollständigen Subtraktion um zwei Items zum Abzählen, ein Item zum Zählen, zwei Items zum Grössenvergleich mit kleinen Zahlen, zwei Items zum Rechnen mit Objektabbildungen und zwei Items zum Lesen von kleinen Zahlen. Die Kompetenzen zum Lösen dieser Aufgaben liegen im Modell von Krajewski (2008) auf den Ebenen 1 (Basisfertigkeiten) und 2 (einfaches Zahlverständnis) und im Modell von Fritz und Ricken (2008) auf den Stufen I (Zählzahl), II (ordinaler Zahlenstrahl) und III (Kardinalität) (vgl. Abschnitt 4.2). Betrachtet man die zehn schwierigsten Items, so wurden diese von durchschnittlich 9.9 % aller Kinder gelöst. Diese Items sind mit Ausnahme des Zahlenlesens der Ebene 3 (tiefes Zahlverständnis) bzw. den Stufen IV (Enthaltensein und Klasseninklusion) und V (Relationalität) zuzuordnen oder sind bereits Aufgaben zum Rechnen. Es handelt sich um vier Items zur Addition auf symbolischer Ebene, um drei Items zum Lesen von grossen Zahlen, um zwei Relationsaufgaben und um ein Item zum Zerlegen einer Zahl. Das schwierigste Item zum ersten Testzeitpunkt war eine Aufgabe zur Addition auf symbolischer Ebene (6 + 8). Dieses Item lösten nur 3.7 % aller Kinder.

Die insgesamt grosse Streuung weist auf grosse Unterschiede in den mathematischen Kompetenzen der Kinder beim ersten Testzeitpunkt hin. Diese Unterschiede konnten teilweise durch Kontextfaktoren erklärt werden, wie die folgenden Analysen zeigen.

8.2.2 Einflussfaktoren auf die mathematische Kompetenz zum ersten Testzeitpunkt

Zur Überprüfung von Zusammenhängen zwischen den interessierenden abhängigen Variablen (mathematische Kompetenzen zu den drei Testzeitpunkten) mit anderen Kontextfaktoren wurden zunächst Korrelationen berechnet.

Tabelle 11: Zusammenhänge zwischen mathematischen Kompetenzen und verschiedenen Kontextfaktoren

	1	2	3	4	5	6	7
1. $MATH^{WLE\_t1}$							
2. $MATH^{WLE\_t2}$	0.82**						
3. $MATH^{WLE\_t3}$	0.74**	0.83**					
4. Kogn. Fähigkeiten	0.51**	0.53**	0.51**				
5. Land	-0.22**	-0.06	-0.15**	-0.14**			
6. Erstsprache	-0.24**	-0.24**	-0.21**	-0.08*	-0.07		
7. Alter t1	0.27**	0.23**	0.19**	0.24**	-0.16**	0.01	
8. Geschlecht	-0.07*	-0.10**	-0.11**	0.09**	-0.01	-0.04	-0.02

Anmerkungen: * p<.05 **p<.01 ***p<.001

Erwartungsgemäss liessen sich hohe positive Korrelationen zwischen den mathematischen Kompetenzen zu den drei Testzeitpunkten und den kognitiven Fähigkeiten beobachten (r zwischen .51 und .83, p< .01), was nach Cohen einem starken Effekt entspricht (Bortz & Schuster, 2016). Die anderen signifikanten Korrelationen waren hingegen eher gering mit schwachen bis mittleren Effekten. Die Personenparameter zwischen den einzelnen Messzeitpunkten korrelierten zwischen r = .74 und r = .83. Diese starken Zusammenhänge weisen darauf hin, dass die relative Position der Kinder über die zwei Kindergartenjahre vergleichsweise stabil war.

Zur Erklärung der Unterschiede in den mathematischen Leistungen zum ersten Testzeitpunkt wurden Mehrebenenanalysen durchgeführt, die die geschachtelte Struktur der Daten mitberücksichtigen (vgl. Abschnitt 7.5.3). Um zu überprüfen, ob die Unterschiede zwischen den Klassen so gross sind, dass eine Mehrebenenanalyse sinnvoll ist, wurde zuerst die Intraklassenkorrelation (ICC) mit dem sogenannten Nullmodell berechnet. In diesem Modell werden keine Prädiktoren eingefügt. Die ICC zeigt, wie gross die Variation zwischen Level-2-Einheiten im Vergleich zur Variation zwischen Level-1-Einheiten ist. Ab einer ICC von 0.10 ist mit verzerrter statistischer Inferenz zu rechnen (Eid, Gollwitzer & Schmitt, 2017). Ab einer ICC von 0.05 sollte die Mehrebene in den Berechnungen berücksichtigt werden (Peugh, 2010). Die Intraklassenkorrelation (ICC) der Mathematikleistung in dieser Studie betrug beim ersten Testzeitpunkt 0.08, beim zweiten 0.07 und beim dritten 0.08. Das heisst, dass zwischen 7 und

8 % der Gesamtvarianz der mathematischen Leistungen durch die Klassenzugehörigkeit erklärt werden kann. Getrennt nach Land betrug die ICC in Deutschland beim ersten Testzeitpunkt 0.11, in der Schweiz 0.07, beim zweiten Testzeitpunkt in Deutschland 0.08, in der Schweiz 0.07 und beim dritten Testzeitpunkt in Deutschland 0.08 und in der Schweiz 0.05. Um zu überprüfen, ob das Random-Intercept-Modell oder das Random-Slope-Modell besser zu den Daten passt, wurde ein Modellvergleich mittels ANOVA berechnet (vgl. Abschnitt 7.5.3). Beim ersten und dritten Testzeitpunkt gab es keine signifikanten Unterschiede zwischen den beiden Modellen ($p > .05$). Allerdings waren jeweils die AIC- und BIC-Werte des Random-Intercept-Modells niedriger. Das bedeutet, dass dieses Modell besser zu den Daten passt. Beim zweiten Testzeitpunkt wurde der Unterschied auch statistisch signifikant ($p < .001$). Auch hier passte das Random-Intercept-Modell besser zu den Daten. In der vorliegenden Untersuchung wurden deshalb die Analysen mit dem Random-Intercept-Modell berichtet. Es wurden alle Variablen in das Modell aufgenommen, die aufgrund der Hypothesen von Interesse waren. Dies waren auf der Individualebene das Alter, das Geschlecht, die Erstsprache und die kognitiven Fähigkeiten. Die kognitiven Leistungen der Kinder wurden zwar erst beim zweiten Testzeitpunkt erhoben (vgl. Abschnitt 7.3.2), doch wurden diese Testergebnisse als Prädiktor bereits beim ersten Testzeitpunkt eingesetzt. Da Intelligenz als ein stabiles Persönlichkeitsmerkmal auch für das Kindergartenalter gilt (Koglin, Janke & Petermann, 2009), kann davon ausgegangen werden, dass die zum zweiten Testzeitpunt gemessenen kognitiven Fähigkeiten auch als valider Prädiktor beim ersten Testzeitpunkt ins Modell aufgenommen werden können. Auf der Klassenebene wurde das Land als Prädiktor eingefügt. Um zu überprüfen, welche Variablen die grösste Erklärungskraft für die Unterschiede liefern, wurde jeweils der erklärende Varianzanteil der Prädiktoren nach Hox et al. (2017) berechnet. Dieser berechnet sich aus der Differenz der Varianz mit und ohne den interessierenden Prädiktor auf der jeweiligen Ebene (vgl. Abschnitt 7.5.3). Tabelle 12 veranschaulicht die Ergebnisse der Mehrebenenanalyse.

Tabelle 12: Einflussfaktoren auf die mathematische Kompetenz zum ersten Testzeitpunkt

	B	*SE*	*p*
Intercept	**-0.79**	0.08	.000
Individualebene			
Alter T1	**0.62**	0.12	.000
Geschlecht	**-0.31**	0.08	.000
Kognitive Fähigkeiten	**0.15**	0.01	.000
Erstsprache	**-0.70**	0.10	.000
Klassenebene			
Land	**0.23**	0.10	.016
	Var	***SD***	
Zufällige Effekte			
Fehlervarianz L1	1.35	1.16	
Fehlervarianz L2	0.05	0.22	
Erklärte Gesamtvarianz (%)	34		

Anmerkungen: Multilevel-Regressionsanalysen: abhängige Variable mathematische Kompetenz zum ersten Testzeitpunkt ($\text{MATH}^{\text{WLE\_t1}}$), N=819. Die Variablen Alter und kognitive Fähigkeiten wurden am Gesamtmittelwert zentriert. Die dichotomen Variablen Geschlecht (0=m, 1=w), Erstsprache (0=Deutsch, 1=andere) und Land (0=CH, 1=D) sind unzentriert.

Insgesamt erklärt das Modell 34 % der Varianz der Kompetenzunterschiede der Kinder beim ersten Testzeitpunkt. Alle eingeführten Prädiktoren waren signifikant. Dabei besassen die allgemeinen kognitiven Fähigkeiten unter Kontrolle der jeweils anderen Variablen die grösste Erklärungskraft für die mathematischen Kompetenzunterschiede zum ersten Testzeitpunkt (B = .15, SE = 0.01, p = .000). Pro Punkt höher in den kognitiven Leistungen erhöhte sich der Wert des Personenparameters ($\text{MATH}^{\text{WLE\_t1}}$) um 0.15 Einheiten. Die erklärte Varianz durch die kognitiven Fähigkeiten betrug 24 %. Die Varianzaufklärung der anderen signifikanten Prädiktoren auf der Individualebene war zwischen 1 (Geschlecht) und 3 % (Alter und Erstsprache) vergleichsweise gering. Je älter die Kinder waren, desto höher waren ihre mathematischen Kompetenzen beim ersten Testzeitpunkt (B = .62, SE = 0.12, p = .000). Zudem wiesen die Jungen höhere Kompetenzen auf als die Mädchen (B = -.31, SE = 0.08, p = .000), ebenso Kinder, deren Erstsprache Deutsch war (B = -.70, SE = 0.01, p = .000). Auf Klassenebene hatte die Zugehörigkeit zum Land einen signifikanten Einfluss und erklärt 17 % der Varianz zwischen den Klassen. Die Kinder in Deutschland zeigten durchschnittlich höhere mathematische Kompetenzen zum ersten Testzeitpunkt als die Kinder in der Schweiz (B = .23, SE = 1.16, p = .000).

8.3 Entwicklung der mathematischen Kompetenzen und ihre Einflussfaktoren

Zur Überprüfung der Hypothesen zur dritten Forschungsfrage (vgl. Abschnitt 7.1.3) wurden Wachstumskurvenmodelle und Mehrebenenanalysen gerechnet. Mittels Wachstumskurvenmodell wurde zuerst die mathematische Entwicklung der Gesamtstichprobe aufgezeigt (Abschnitt 8.3.1). Um den Einfluss verschiedener Faktoren auf die Kompetenzentwicklung der Kinder zu überprüfen, erfolgten danach Mehrebenenanalysen. Die Ergebnisse der Einflussfaktoren auf die mathematische Entwicklung vom ersten zum zweiten Testzeitpunkt werden in Abschnitt 8.3.2 berichtet, jene zur Entwicklung vom zweiten zum dritten Testzeitpunkt in Abschnitt 8.3.3.

8.3.1 Entwicklungsverlauf der Gesamtstichprobe

Die folgende Darstellung veranschaulicht die Mittelwerte und die Standardabweichung über die drei Testzeitpunkte.

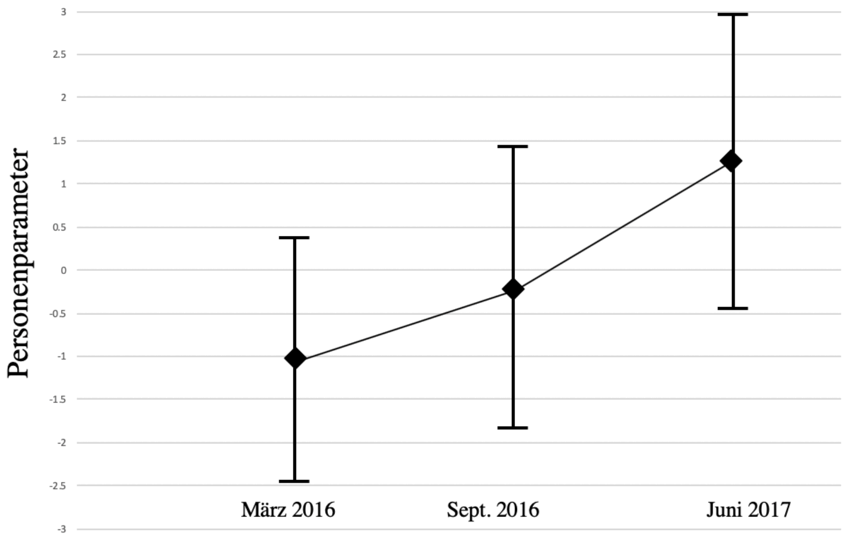

Abbildung 5: Mittelwerte und Standardabweichungen über die drei Testzeitpunkte

Der Mittelwert der Personenfähigkeitsparameter der Kinder betrug beim ersten Testzeitpunkt $M = -1.02$ und die Standardabweichung $SD = 1.45$. Beim zweiten Testzeitpunkt lag der Mittelwert bei $M = -.22$ und die Standardabweichung bei $SD = 1.64$ und beim dritten Testzeitpunkt ergaben sich $M = 1.27$ und $SD = 1.69$. Der Mittelwert hat sich sowohl vom ersten zum zweiten als auch vom zweiten zum dritten Testzeitpunkt erhöht. Auch die Standardabweichung vergrösserte sich vom ersten zum zweiten Testzeitpunkt und auch noch vom zweiten zum dritten, wenn auch wesentlich geringer. Diese Zunahme an mathematischer Kompetenz der Gesamtstichprobe zeigte sich auch in der Verteilung der Personen- und Itemparameter über die drei Testzeitpunkte, wie folgende graphische Darstellung verdeutlicht.

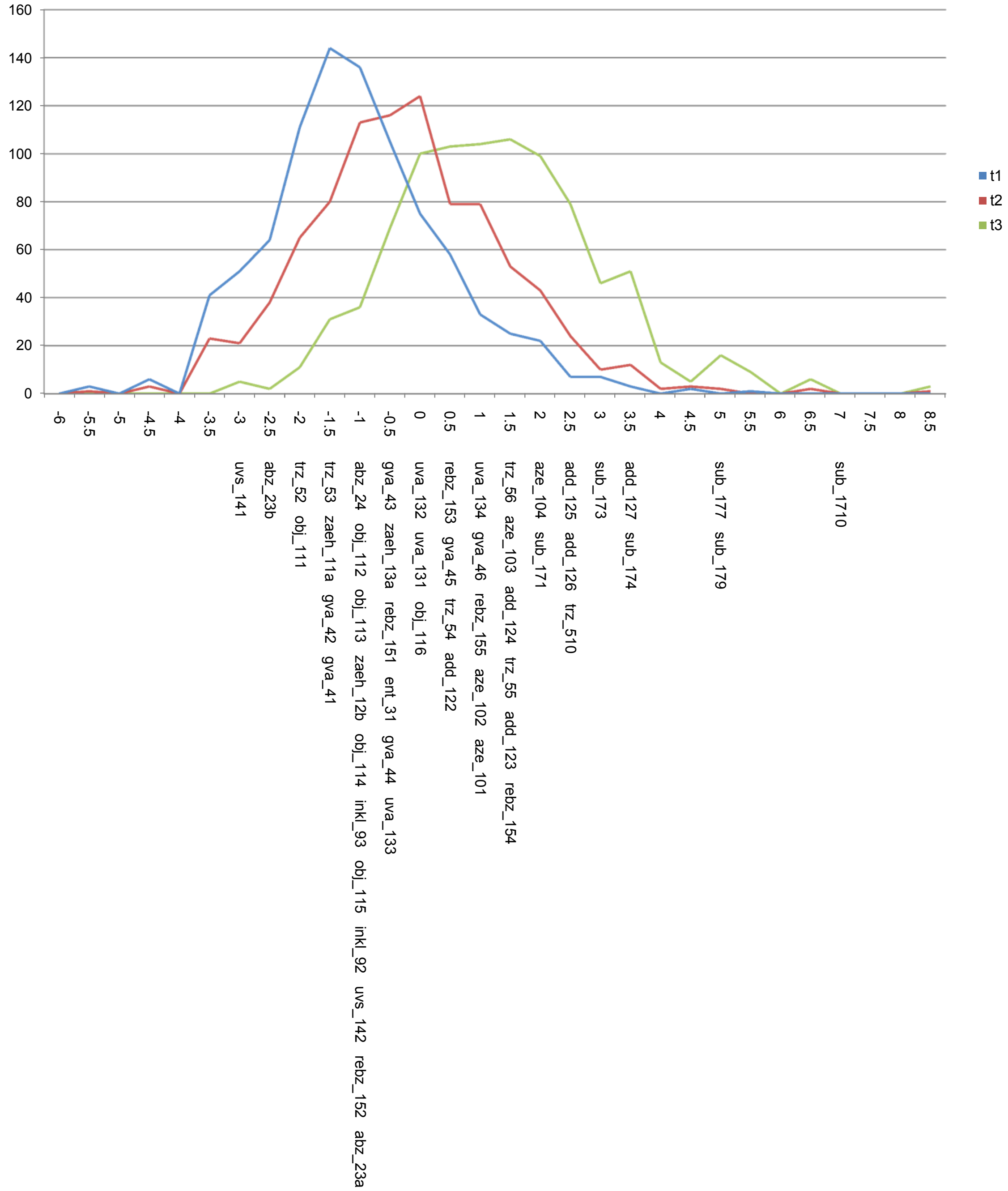

Abbildung 6: Personen- und Itemparameter über die drei Testzeitpunkte

In der Abbildung sind die drei Verteilungen (t1 bis t3) der Kinder (Anzahl auf y-Achse) nach Fähigkeit bzw. nach Schwierigkeit der Items (x-Achse) zu erkennen. Die Personen- und die Itemparameter wurden zu Einheiten von jeweils 0.5 Logits zusammengefasst. Am schwierigsten waren drei Items zur Subtraktion, am einfachsten erwies sich ein Item zur unvollständigen Subtraktion (bildgestützt) und eines zum Zählen von Objekten (Wie viele Schildkröten sind alle zusammen?). Die Kurven über die drei Testzeitpunkte haben sich von links nach rechts verschoben. Darin spiegelt sich die Zunahme der Personenfähigkeit bzw. der Kompetenzerwerb der Kinder vom ersten zum zweiten und vom zweiten zum dritten Messzeitpunkt wider. Folgend wurde dieser Entwicklungsverlauf der Gesamtstichprobe statistisch analysiert. Dazu wurde ein latentes Wachstumskurvenmodell gemäss folgendem Modell berechnet.

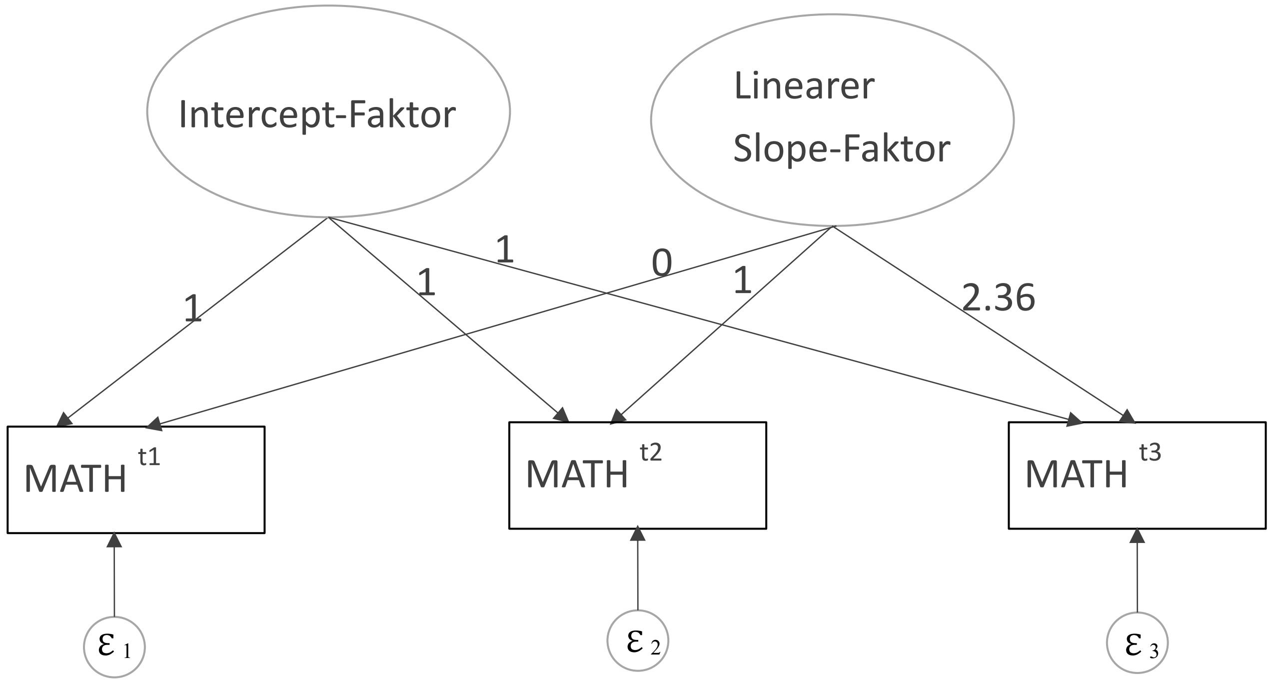

Abbildung 7: Wachstumskurvenmodell mit fixierten Slope-Faktoren

Das Modell legt dar, dass gemäss den theoretischen Annahmen alle erhobenen mathematischen Kompetenzen durch den latenten Intercept-Faktor beeinflusst werden. Dieser Faktor repräsentiert interindividuelle Unterschiede in der Ausgangsleistung, das heisst, die mathematische Kompetenz zum ersten Testzeitpunkt ($MATH^{t1}$). Entsprechend wurden die Ladungen dieses Intercept-Faktors auf die Kompetenzen zu den drei Testzeitpunkten auf 1 fixiert (Eingangsleistung bleibt stabil). Der Slope-Faktor repräsentiert die Veränderung über die Zeit. Es wurde von einem linearen Wachstum ausgegangen. Wenn die Abstände zwischen den zeitlichen Messungen gleich wären, würden die Ladungen auf 0 (kein Wachstum zum ersten Testzeitpunkt), auf 1 und auf 2 für die weiteren Messzeitpunkte fixiert. In vorliegender

Studie gab es allerdings eine Differenz in den zeitlichen Abständen zwischen den Testzeitpunkten. Zwischen dem ersten und zweiten Testzeitpunkt lagen 28 Wochen. Zwischen dem zweiten und dritten Testzeitpunkt waren es 38 Wochen. Dabei wurde jeweils die mittlere Woche der zwei Testphasen als Referenzpunkt genommen. Dies ergab ein Verhältnis von 1.36. Entsprechend wurde der dritte Faktor nicht auf 2, sondern auf 2.36 fixiert. Nicht berücksichtigt werden konnte indes die effektive Zeit, die die Kinder im Kindergarten verbracht haben, da die Kindergartenzeiten und die Dauer der Ferien sich sowohl in den in den verschiedenen Kantonen der Schweiz als auch in Deutschland unterscheiden. Tabelle 13 fasst die Ergebnisse der Analyse mit dem Wachstumskurvenmodell zusammen.

Tabelle 13: Ergebnisse LGCM mit fixiertem Slope-Faktor zum dritten Testzeitpunkt

	Intercept		Slope	
	B	*SE*	*B*	*SE*
WLE_t1	1.00		0.00	
WLE_t2	1.00		1.00	
WLE_t3	1.00		2.36	
Mittelwert	**-1.03***	.05	**.98***	.02
Varianz	**2.03***	.12	**.25***	.03
slope mit intercept		*B* = **-.09***	*SD* = .04	
CFI/RMSEA/SRMR	0.98/0.21/0.04			

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Im Ergebnis wurde deutlich, dass der Modellfit nicht überall befriedigend war ($CFI = 0.98$; $RMSEA = 0.21$; $SRMR = 0.04$). Insbesondere der RMSEA-Wert war sehr hoch (vgl. Abschnitt 7.5.5). Deshalb wurde ein zweites Modell gerechnet, bei dem der dritte Slope-Faktor frei geschätzt wurde. Tabelle 14 zeigt die Ergebnisse dieses Modells.

Tabelle 14: Ergebnisse LGCM mit frei geschätztem Slope-Faktor zum dritten Testzeitpunkt

	Intercept		Slope	
	B	*SE*	*B*	*SE*
WLE_t1	1.00		0.00	
WLE_t2	1.00		1.00	
WLE_t3	1.00		2.86	.09
Mittelwert	**-1.02*****	.05	**.80*****	.03
Varianz	**2.01*****	.11	**.20*****	.03
slope mit intercept		*B* = **-.07***	*SE* = .03	
CFI/RMSEA/SRMR		1.00/0.00/0.00		

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Der Modellfit war nun deutlich besser (*CFI* = 1.00; *RMSEA* = 0.00; *SRMR* = 0.00). Zudem war der geschätzte Slope-Faktor zum dritten Testzeitpunkt mit 2.86 deutlich höher als der aufgrund der zeitlichen Differenz zwischen den Messungen angenommene Wert von 2.36. Dies gibt einen Hinweis darauf, dass der Kompetenzzuwachs zwischen dem zweiten und dritten Testzeitpunkt grösser war als zwischen dem ersten und zweiten Testzeitpunkt. Um zu überprüfen, ob der Kompetenzzuwachs zwischen dem zweiten und dritten Testzeitpunkt (Spieleeinsatz) tatsächlich höher war, wurde ein zweiter Slope-Faktor „Intervention“ gemäss untenstehendem Modell aufgenommen. Damit kann überprüft werden, ob der Unterschied zwischen dem aufgrund der zeitlichen Differenz errechneten Slope-Faktor (2.36) und dem frei geschätzten (2.86) statistisch signifikant ist.

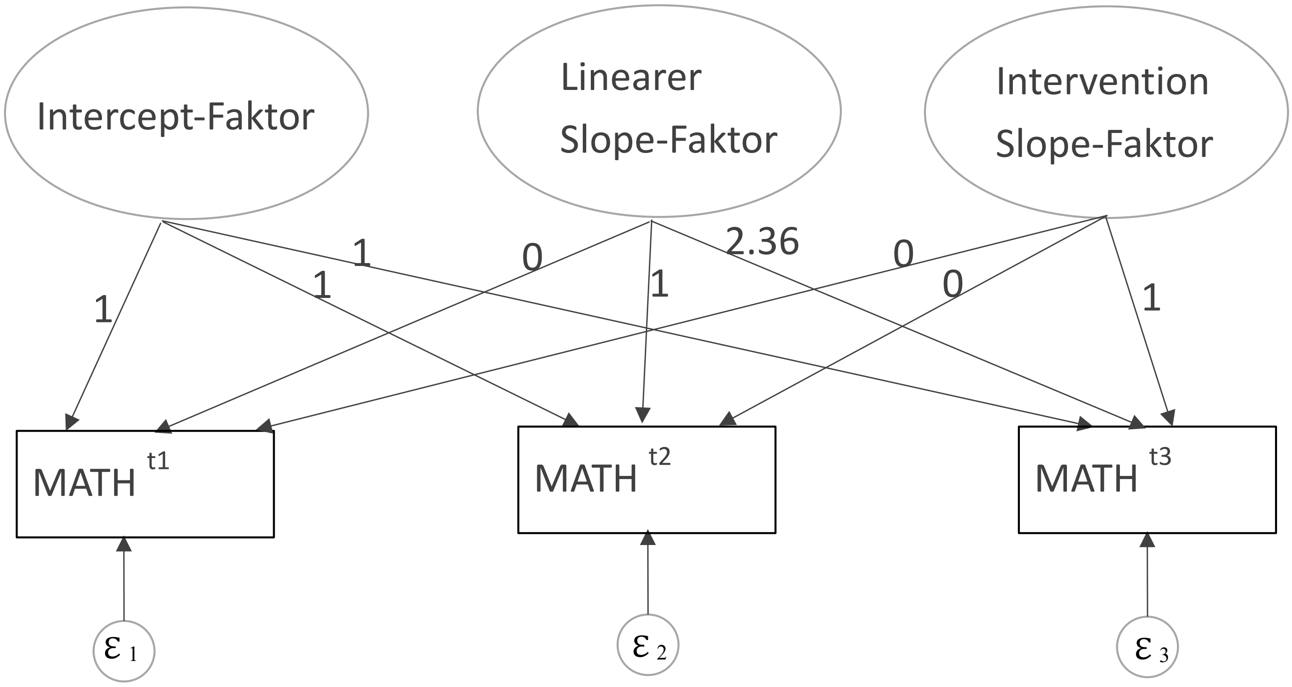

Abbildung 8: Wachstumskurvenmodell mit Slope-Faktor Intervention

Tabelle 15 veranschaulicht die Ergebnisse der Analyse:

Tabelle 15: Ergebnisse LGCM mit Interventions-Slope-Faktor

	Intercept		Slope		Intervention	
	B	*SD*	*B*	*SD*	*B*	*SB*
WLE_t1	1.00		0.00		0.00	
WLE_t2	1.00		1.00		0.00	
WLE_t3	1.00		2.36		1.00	
Mittelwert	**-1.02*** **	.05	**.80*** **	.03	**.40*** **	.06
Varianz	1.97		**.23** **	.07	.00	
CFI/RMSEA/SRMR			1.00/0.01/0.00			

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Das Modell zeigte insgesamt gute Fit-Werte (*CFI* =1.00; *RMSEA* = 0.01; *SRMR* = 0.00). Der Unterschied zwischen dem Slope-Faktor (2.36) und dem frei geschätzten (2.86) war statistisch signifikant ($B = .40$, $p < 0.001$). Das bedeutet, dass die Kinder zwischen dem zweiten und dritten Testzeitpunkt unter Berücksichtigung der zeitbezogenen Unterschiede einen grösseren mathematischen Kompetenzzuwachs erzielten als zwischen dem ersten und zweiten.

Die Ergebnisse des Wachstumskurvenmodells mit Intercept- und Slope-Faktor (vgl. Tabelle 13) liessen darüber hinaus auch erkennen, dass die Ausgangsleistung mit dem Leistungszuwachs signifikant negativ korrelierte ($B = -.07$, $p < 0.05$). Das bedeutet, dass Kinder mit niedriger Ausgangsleistung einen grösseren Kompetenzzuwachs hatten. Um zu überprüfen, ob dies auch unter Kontrolle der Kontextfaktoren der Fall ist, wurden diese Variablen als Kovariate ins Modell aufgenommen.

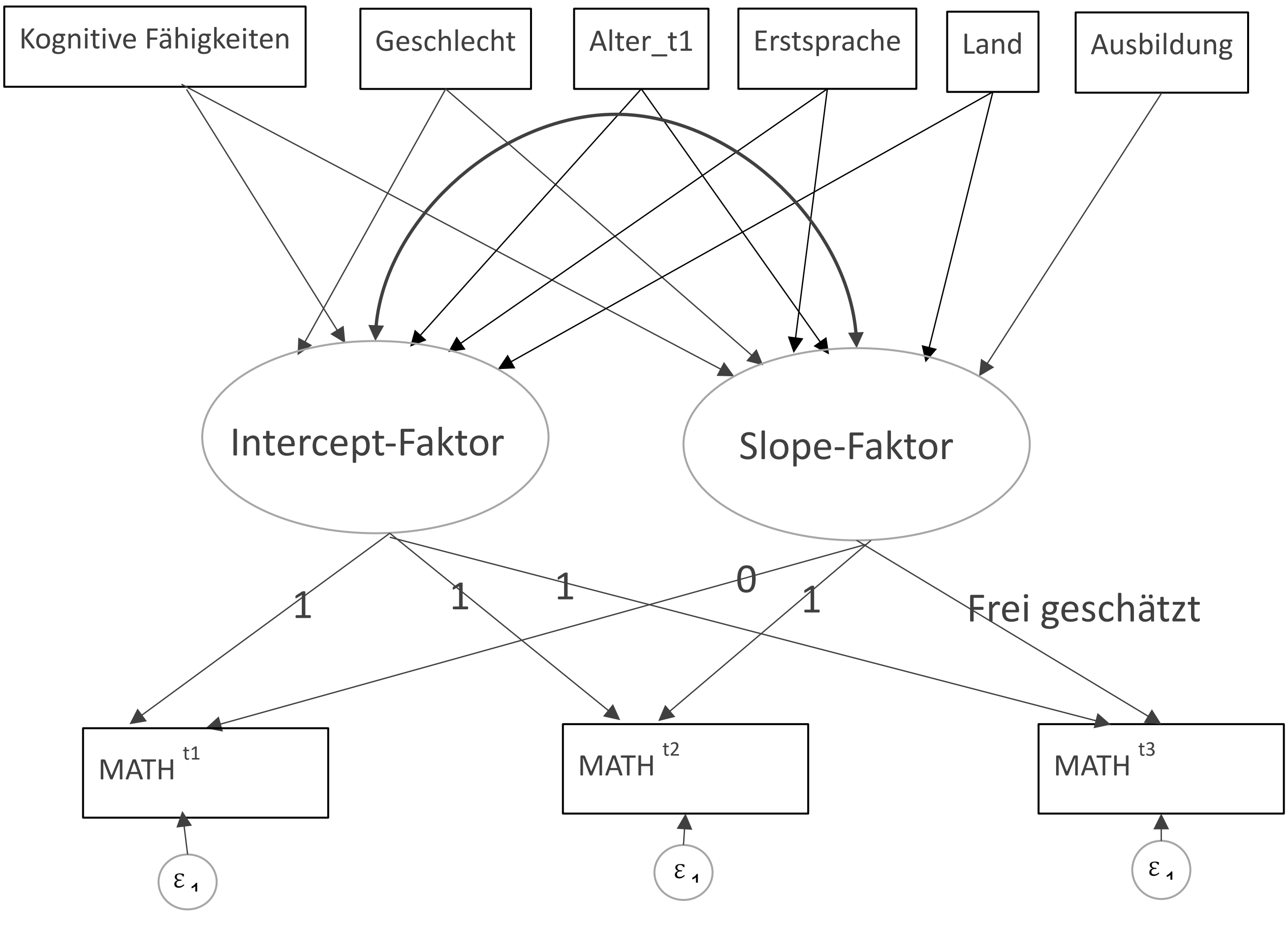

Abbildung 9: Wachstumskurvenmodell mit Prädiktoren

Tabelle 16 zeigt die Ergebnisse der beiden Analysen im Vergleich.

Tabelle 16: Ergebnisse Wachstumskurvenmodell mit und ohne Kovariaten

	Modell ohne Kovariaten				Modell mit Kovariaten			
	Intercept		Slope		Intercept		Slope	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Mittelwert	**-1.02*****	.05	**.80*****	.03	**-5.46*****	.58	**1.51*****	.19
Varianz	**2.01*****	.11	**.20*****	.03	**1.17*****	.07	**.13*****	.02
slope mit intercept	$B =$ **-.07*** $SE = .03$ $p = 0.01$				$B = -0.04$ $SE = 0.02$ $p = 0.07$			
CFI/RMSEA/SRMR	1.00/0.00/0.00				0.99/0.06/0.02			

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Das Modell mit den Kovariaten wies ebenfalls gute Fit-Werte auf ($CFI = 0.99$; $RMSEA = 0.06$; $SRMR = 0.02$). Es wurde nun ersichtlich, dass unter Kontrolle der Variablen kognitive Fähigkeiten, Geschlecht, Alter, Erstsprache, Land und Ausbildung der Fachkräfte die Korrelation zwischen mathematischen Leistungen zum ersten Testzeitpunkt und dem Kompetenzzuwachs knapp nicht signifikant ($B = -0.04$, $p = 0.07$) wurde. In der Tendenz verzeichneten aber auch hier Kinder mit niedrigen mathematischen Kompetenzen einen höheren Leistungszuwachs. Es fällt auf, dass der mittlere Ausgangswert (Intercept) im Modell mit Kovariaten deutlich niedrig (-5.46) ist als ohne Kovariaten (-1.02). Das heisst, dass die mathematischen Leistungen beim ersten Testzeitpunkt unter Kontrolle der eingefügten Prädiktoren niedriger sind. Dies lässt sich vermutlich mit der hohen Korrelation von kognitiven Fähigkeiten und mathematischen Kompetenzen beim ersten Testzeitpunkt ($r = .51$) erklären.

Des Weiteren liess sich beobachten, dass die Kinder insgesamt einen signifikanten Leistungszuwachs hatten ($B = 1.51$, $p < 0.001$) und dass sich die Kinder in diesem Leistungszuwachs signifikant unterschieden ($B = 0.13$, $p < 0.001$). Welchen Einfluss Kontextfaktoren auf diese Unterschiede in der Leistungsentwicklung haben, wird im Folgenden untersucht.

8.3.2 Einflussfaktoren auf den mathematischen Kompetenzzuwachs zwischen dem ersten und zweiten Testzeitpunkt

Sechs Monate nach der ersten Erhebung wurden die mathematischen Leistungen der Kinder zum zweiten Mal, wiederum mit dem TEDI-MATH (vgl. Abschnitt 7.3.1), gemessen. Die Kinder befanden sich nun im ersten Quartal des letzten Kindergartenjahres. Der Mittelwert der Personenparameter (WLE) der Untersuchungsstichprobe (N = 894) betrug M = -0.22 mit einer Standardabweichung von SD = 1.64.

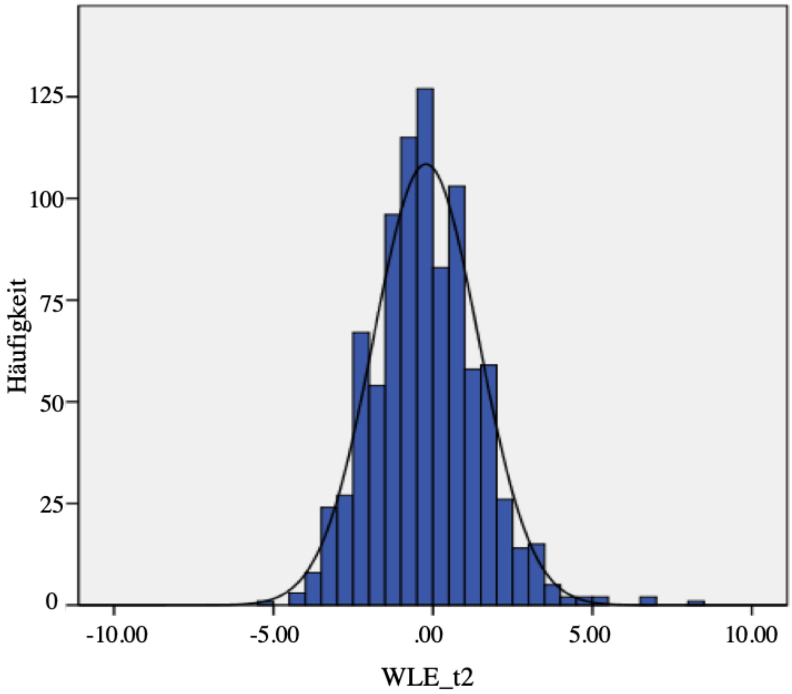

Abbildung 10: Mathematische Kompetenzen zum zweiten Testzeitpunkt

Zur Erklärung von Unterschieden in der Kompetenzentwicklung wurde nun das mathematische Vorwissen beim ersten Testzeitpunkt als Prädiktor ins Modell der Mehrebenenanalyse aufgenommen und somit kontrolliert. Die aufgrund der Hypothesen interessierenden Kontextfaktoren wurden wieder als Prädiktoren eingeführt und erklären damit Unterschiede im Zuwachs der mathematischen Leistungen der Kinder zwischen dem ersten und zweiten Testzeitpunkt. Es waren dies das Alter, das Geschlecht, die kognitiven Fähigkeiten und die Erstsprache auf der Individualebene. Auf der Klassenebene wurden das Land und zusätzlich die Ausbildung der pädagogischen Fachkräfte aufgenommen. Wie in Abschnitt 2.1.2 beschrieben, unterscheiden sich die Fachpersonen in den beiden Länder bezüglich der Ausbildung zur pädagogischen Fachkraft im Kindergarten erheblich. Insbesondere ist der Anteil von Fachkräften mit einer akademischen Ausbildung in der Schweiz mit 46.1 % deutlich höher als in Deutschland mit 14.6 %. Die Variable Ausbildung gibt an, ob die akademische Ausbildung verglichen mit der nicht akademischen Ausbildung einen Einfluss auf die Kompetenzentwicklung der Kinder hatte. Zusätzlich wurde überprüft, ob die Interaktion von Ausbildung und Land die Kompetenzentwicklung beeinflusste. Um zu überprüfen, welche Variablen die grösste Erklärungskraft für die Unterschiede im Kompetenzzuwachs liefern, wurde wiederum der erklärende Varianzanteil der jeweiligen Prädiktoren nach Hox et al. (2017) berechnet (vgl. Abschnitt 7.5.3). Tabelle 17 gibt einen Überblick über die Ergebnisse der Mehrebenenanalyse.

Tabelle 17: Einflussfaktoren auf die mathematische Kompetenzentwicklung zwischen dem ersten und zweiten Testzeitpunkt

	B	***SE***	***p***
Intercept	-0.01	0.08	.939
Individualebene			
MATH^{WLE_t1}	**0.81**	0.03	.000
Alter T1	-0.04	0.09	.640
Geschlecht	**-0.23**	0.06	.000
Kognitive Fähigkeiten	**0.06**	0.01	.000
Erstsprache	**-0.26**	0.08	.002
Klassenebene			
Land	0.10	0.18	.567
Ausbildung	-0.53	0.10	.598
Land*Ausbildung	-0.18	0.20	.367
	Var	***SD***	
Zufällige Effekte			
Fehlervarianz L1	0.78	0.88	
Fehlervarianz L2	0.07	0.26	
Erklärte Gesamtvarianz (%)	69		

Anmerkungen: Multilevel-Regressionsanalysen: abhängige Variable numerische Kompetenz zum zweiten Testzeitpunkt (MATH^{WLE_t2}), N=819. Die Variablen MATH^{WLE_t1} , Alter und kognitive Fähigkeiten wurden am Gesamtmittelwert zentriert. Die dichotomen Variablen Geschlecht (0=m, 1=w), Erstsprache (0=Deutsch, 1=andere), Land (0=CH, 1=D) und Ausbildung (0=akademisch, 1 = nicht akademisch) sind unzentriert.

Insgesamt erklärte das Modell 69 % der Varianz in den mathematischen Kompetenzen der Kinder zum zweiten Testzeitpunkt. Dabei besass das Vorwissen der Kinder zum ersten Testzeitpunkt erwartungsgemäss die grösste Erklärungskraft (B = .81, SE = 0.03, p = .000); es erklärte 55 % der Varianz in den mathematischen Kompetenzen der Kinder zum zweiten Testzeitpunkt. Als weitere signifikante Prädiktoren auf der Individualebene erwiesen sich das Geschlecht (B = -.21, SE = 0.06, p = .000), die kognitiven Fähigkeiten (B = 0.06, SE = 0.01,

p = .000) und die Erstsprache (B = -0.26, SE = 0.08, p = .002). Je höher die kognitiven Fähigkeiten waren, desto grösser waren auch die Fortschritte der Kinder zwischen dem ersten und zweiten Testzeitpunkt. Die Jungen erzielten dabei einen grösseren Kompetenzzuwachs, ebenso wie die Kinder mit Erstsprache Deutsch. Die Varianzaufklärung dieser Prädiktoren war aber eher gering und reichte von weniger als 1 % (Erstsprache) über 3 % (Geschlecht) bis 7 % (kognitive Fähigkeiten). Keinen Einfluss auf die Kompetenzentwicklung zwischen dem ersten und zweiten Testzeitpunkt hatte hingegen das Alter. Auf Klassenebene liess sich hinsichtlich des Landes kein signifikanter Einfluss ermitteln und auch die Ausbildung der pädagogischen Fachkräfte (akademisch versus nicht akademisch) sowie die Interaktion von Ausbildung und Land beeinflussten die mathematische Leistungsentwicklung nicht.

8.3.3 Einflussfaktoren auf den mathematischen Kompetenzzuwachs zwischen dem zweiten und dritten Testzeitpunkt

Die folgenden Analysen beziehen sich auf den Zeitraum von September (t2) bis Juni (t3). Die Kinder befanden sich beim dritten Testzeitpunkt kurz vor dem Eintritt in die Grundschule. Die Leistungen der Kinder wurden erneut mit dem TEDI-MATH (vgl. Abschnitt 7.3.1) gemessen. Der Mittelwert der Personenparameter (WLE) der Untersuchungsstichprobe (N = 894) betrug M = 1.27 mit einer Standardabweichung von SD = 1.69.

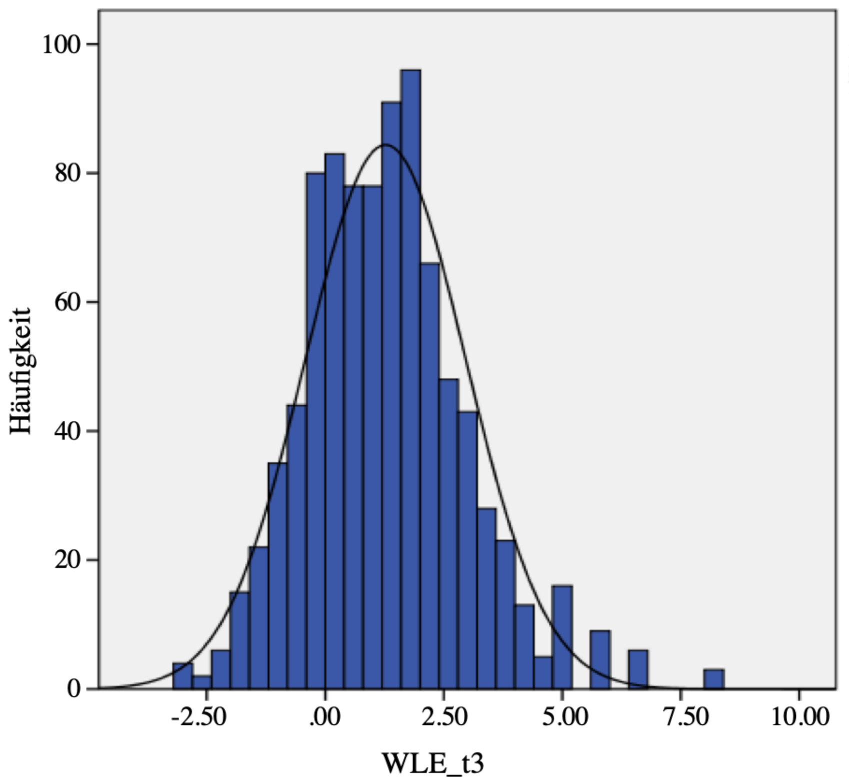

Abbildung 11: Mathematische Kompetenzen zum dritten Testzeitpunkt

Zur Erklärung von Unterschieden in der Kompetenzentwicklung zwischen dem zweiten und dritten Testzeitpunkt wurde nun das mathematische Vorwissen beim zweiten Testzeitpunkt als Prädiktor ins Modell aufgenommen und somit kontrolliert. Die aufgrund der Hypothesen interessierenden Kontextfaktoren wurden wieder als Prädiktoren eingeführt und erklären damit Unterschiede im Zuwachs der mathematischen Leistungen der Kinder zwischen dem zweiten und dritten Testzeitpunkt. Es wurde der Einfluss der kognitiven Fähigkeiten, des Alters, der Erstsprache, des Landes und der Ausbildung der pädagogischen Fachkräfte auf die Kompetenzentwicklung untersucht. Um zu überprüfen, welche Variablen die grösste Erklärungskraft für die Unterschiede liefern, wurde wiederum der erklärende Varianzanteil der interessierenden Prädiktoren nach Hox et al. (2017) berechnet (vgl. Abschnitt 7.5.3). Tabelle 18 enthält die Ergebnisse der Mehrebenenanalyse beim dritten Testzeitpunkt.

Tabelle 18: Einflussfaktoren auf die mathematische Kompetenzentwicklung zwischen dem zweiten und dritten Testzeitpunkt

	B	***SE***	***p***
Intercept	**1.48**	0.09	.000
Individualebene			
$MATH^{WLE\_t2}$	**0.81**	0.02	.000
Alter T1	-0.16	0.09	.077
Geschlecht	**-0.14**	0.06	.027
Kognitive Fähigkeiten	**0.03**	0.01	.000
Erstsprache	-0.07	0.08	.397
Klassenebene			
Land	**-0.48**	0.18	.008
Ausbildung	-0.07	0.10	.483
Land*Ausbildung	-0.12	0.20	.557
	Var	***SD***	
Zufällige Effekte			
Fehlervarianz L1	0.77	0.88	
Fehlervarianz L2	0.07	0.27	
Erklärte Gesamtvarianz (%)	71		

Anmerkungen: Multilevel-Regressionsanalysen: abhängige Variable numerische Kompetenz zum dritten Testzeitpunkt ($MATH^{WLE\_t3}$), N=819. Die Variablen $MATH^{WLE\_t2}$, Alter und kognitive Fähigkeiten wurden am Gesamtmittelwert zentriert. Die dichotomen Variablen Geschlecht (0=m, 1=w), Erstsprache (0=Deutsch, 1=andere), Land (0=CH, 1=D) und Ausbildung (0=akademisch, 1 = nicht akademisch) sind unzentriert.

Insgesamt erklärte das Modell 71 % der Varianz der Unterschiede in den mathematischen Kompetenzen beim dritten Testzeitpunkt. Das mathematische Wissen beim zweiten

Testzeitpunkt besass mit 60 % Varianzaufklärung wiederum die grösste Erklärungskraft ($B = .81$, $SE = 0.02$, $p = .000$). Auch zwischen dem zweiten und dritten Testzeitpunkt machten Kinder mit höheren kognitiven Fähigkeiten grössere Fortschritte. Die kognitiven Fähigkeiten erklärten 4 % der Varianz im Kompetenzzuwachs ($B = .03$, $SE = 0.01$, $p = .000$). Das Geschlecht war ebenfalls ein signifikanter Prädiktor, hatte aber nur einen Varianzanteil von knapp 1 % ($B = -.14$, $SE = -0.06$, $p = 0.027$). Bei den Jungen war wiederum ein grösserer Kompetenzzuwachs zu verzeichnen. Keinen Einfluss auf die Kompetenzentwicklung zwischen dem zweiten und dritten Testzeitpunkt hatte hingegen das Alter ($B = -0.16$, $SE = 0.09$, $p = .077$), wobei tendenziell die jüngeren Kinder einen höheren Leistungszuwachs erzielten. Die Erstsprache ($B = -.07$, $SE = 0.08$, $p = .397$) hatte im Gegensatz zu t2 keinen Einfluss mehr. Auf Klassenebene hatte das Land beim dritten Testzeitpunkt einen signifikanten Einfluss auf die Leistungsentwicklung ($\beta = -0.48$, $SE = 0.18$, $p = 0.008$) mit einer Varianzaufklärung von 22 %. Kinder in der Schweiz wiesen einen grösseren mathematischen Kompetenzzuwachs als Kinder in Deutschland auf. Die Ausbildung der pädagogischen Fachkräfte (akademisch versus nicht akademisch) beeinflusste den Kompetenzzuwachs indes nicht; das gleiche gilt für die Interaktion zwischen Ausbildung und Land.

8.4 Überblick zum Entwicklungsverlauf der verschiedenen Gruppen

Nachfolgend werden die Entwicklungsverläufe von Mädchen und Jungen, von Kindern mit Erstsprache Deutsch und einer anderen Erstsprache sowie von Kindern aus Deutschland und der Schweiz beschrieben und graphisch dargestellt.

8.4.1 Entwicklungsverlauf von Mädchen und von Jungen

Die Mehrebenenanalysen haben ergeben, dass die Jungen mit höheren Kompetenzen starteten. Zum ersten Testzeitpunkt betrug der Mittelwert der Personenparameter der Jungen $M = -.93$ ($SD = 1.58$) und der der Mädchen $M = -1.12$ ($SD = 1.30$). Auch der Kompetenzzuwachs der Jungen war zwischen dem ersten und dem zweiten sowie zwischen dem zweiten und dem dritten Testzeitpunkt grösser. Zum zweiten Testzeitpunkt betrug der Mittelwert der Jungen $M = -.06$ ($SD = 1.80$) und der der Mädchen $M = -.38$ ($SD = 1.45$). Zum dritten Testzeitpunkt betrug der Mittelwert der Personenparameter der Jungen $M = 1.45$ ($SD = 1.84$) und derjenige der Mädchen $M = 1.09$ ($SD = 1.50$). Erwartungsgemäss waren die Mittelwertunterschiede beim zweiten und dritten Testzeitpunkt statistisch bedeutsam (F [3,818] = 149.288, $p = .000$) und (F [3,844] = 113.501, $p = .000$). Als Kovariate wurden in den Varianzanalysen jeweils die Variablen eingefügt, die sich in den Mehrebenenanalysen als signifikant herausstellten. Die folgende Graphik veranschaulicht die Kompetenzentwicklung von Mädchen und Jungen über die drei Testzeitpunkte hinweg.

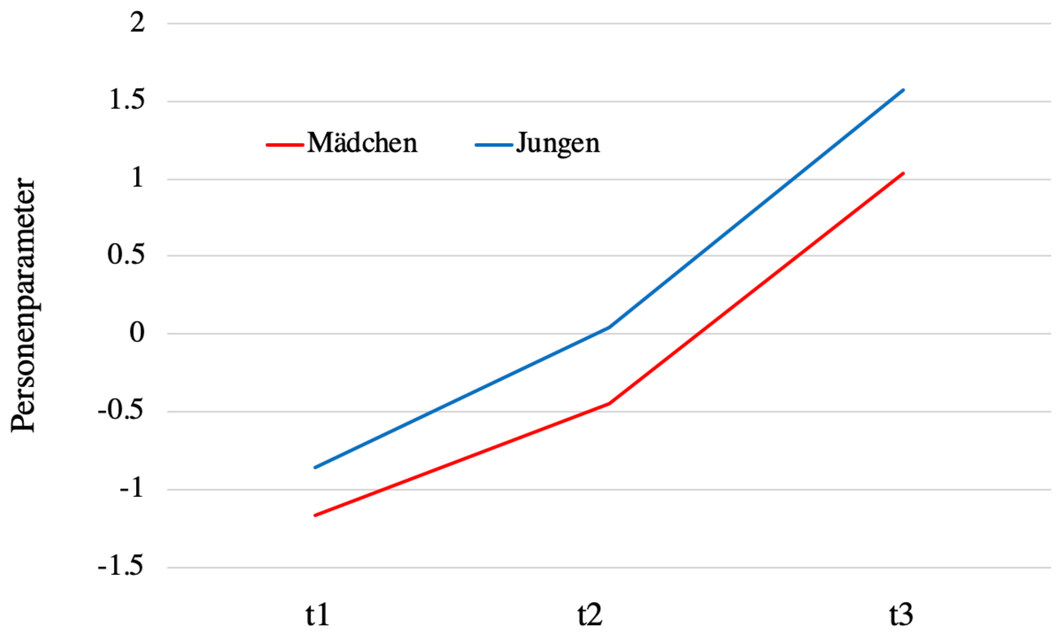

Abbildung 12: Mathematische Kompetenzentwicklung getrennt nach Geschlecht

Wie in der Abbildung zu erkennen, starteten die Jungen mit höheren mathematischen Kompetenzen und der Kompetenzvorsprung vergrösserte sich über die Zeitpunkte noch leicht.

8.4.2 Entwicklungsverlauf von Kindern mit Erstsprache Deutsch und von Kindern mit einer anderen Erstsprache

Bezüglich des Faktors der Erstsprache ging aus den Mehrebenenanalysen hervor, dass diese einen Einfluss auf die mathematischen Kompetenzen zum ersten Testzeitpunkt hatte. Der Mittelwert der Kinder mit Erstsprache Deutsch betrug beim ersten Testzeitpunkt $M = -0.85$ ($SD = 1.41$), derjenige der Kinder mit einer anderen Erstsprache $M = -1.63$ ($SD = 1.42$). Die Erstsprache beeinflusste jedoch auch die Leistungsentwicklung vom ersten zum zweiten Testzeitpunkt. Der Mittelwert der Kinder mit Erstsprache Deutsch betrug beim zweiten

Testzeitpunkt M = -0.00 (SD = 1.59), derjenige der Kinder mit einer anderen Erstsprache M = -.95 (SD = 1.63). Diese Differenz ist statistisch signifikant (F[3,818] = 149.288, p = .000). Auf die Leistungsentwicklung zwischen dem zweiten und dritten Testzeitpunkt hatte die Erstsprache hingegen keinen Einfluss mehr. Trotzdem unterschieden sich die Kinder mit unterschiedlicher Erstsprache bezüglich der Mittelwerte auch beim dritten Testzeitpunkt signifikant (F[4,817] = 96.153, p = .000). Der Mittelwert der Kinder mit Erstsprache Deutsch betrug beim dritten Testzeitpunkt M = 1.47 (SD = 1.63), derjenige der Kinder mit einer anderen Erstsprache M = .62 (SD = 1.74). Die folgende Graphik veranschaulicht die mathematische Kompetenzentwicklung von Kindern mit Erstsprache Deutsch und einer anderen Erstsprache über die drei Testzeitpunkte.

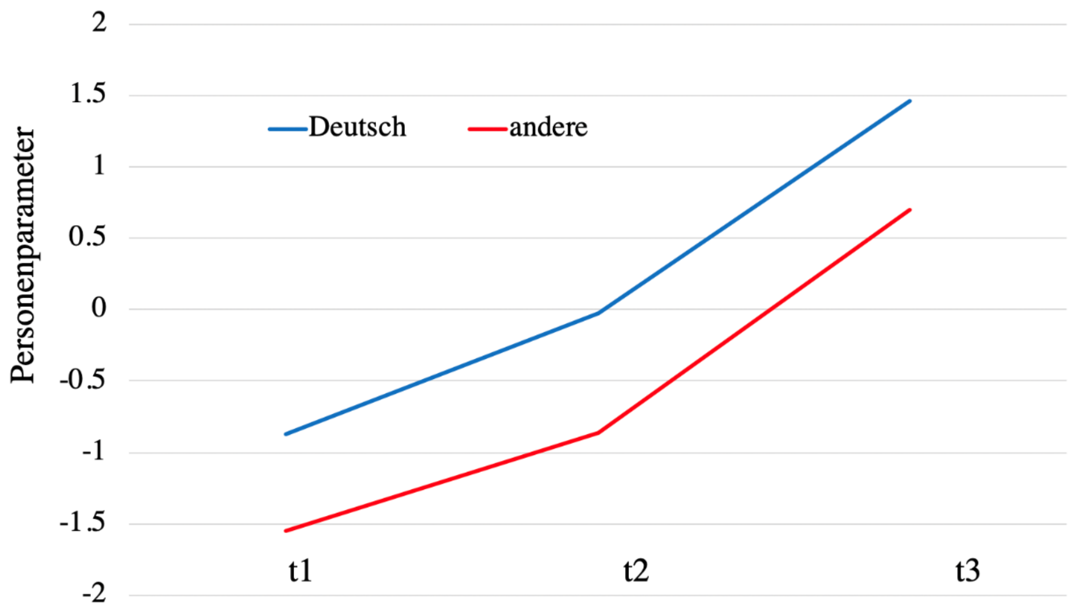

Abbildung 13: Mathematische Kompetenzentwicklung getrennt nach Erstsprache

Die Abbildung verdeutlicht, dass Kinder mit einer anderen Erstsprache als Deutsch mit niedrigeren mathematischen Kompetenzen starteten und dass sich der Kompetenzrückstand vom ersten zum zweiten Testzeitpunkt noch leicht vergrösserte. Zwischen dem zweiten und dem dritten Testzeitpunkt verlief die mathematische Kompetenzentwicklung von Kindern mit Erstsprache Deutsch und einer anderen Erstsprache parallel.

8.4.3 Entwicklungsverlauf von Kindern aus Deutschland und von Kindern aus der Schweiz

Die Kinder in Deutschland erzielten beim ersten Testzeitpunkt unter Kontrolle individueller Faktoren wie Geschlecht, Alter, Erstsprache und kognitive Fähigkeiten höhere Leistungen als die Kinder in der Schweiz. Der Mittelwert der Kinder in Deutschland betrug $M = -1.05$ ($SD = 1.33$), derjenige der Kinder in der Schweiz $M = -1.00$ ($SD = 1.53$). Zwischen dem ersten und dem zweiten Testzeitpunkt entwickelten sich die Kinder in beiden Ländern gleich. Der Mittelwert der Kinder in Deutschland betrug zum zweiten Testzeitpunkt $M = -0.33$ ($SD = 1.55$), derjenige der Kinder in der Schweiz $M = -0.14$ ($SD = 1.70$). Die Unterschiede in den mittleren mathematischen Kompetenzen waren unter Kontrolle der signifikanten Prädiktoren (Geschlecht, Erstsprache, kognitive Fähigkeiten) statistisch aber signifikant ($F[4,817] = 111.989$, $p = .000$). Die Kinder in Deutschland wiesen hier noch höhere mathematische Kompetenzen auf. Vom zweiten zum dritten Testzeitpunkt erzielten jedoch die Kinder in der Schweiz einen grösseren Kompetenzzuwachs. Der Mittelwert der Kinder in Deutschland betrug zum dritten Testzeitpunkt $M = 0.90$ ($SD = 1.58$), derjenige der Kinder in der Schweiz $M = 1.48$ ($SD = 1.73$). Dieser Unterschied in den Kompetenzen beim dritten Testzeitpunkt war nun zugunsten der Kinder in der Schweiz signifikant ($F[3,844] = 113.509$, $p = .000$). Die folgende Graphik veranschaulicht die Kompetenzentwicklung von Kindern in Deutschland und von Kindern in der Schweiz.

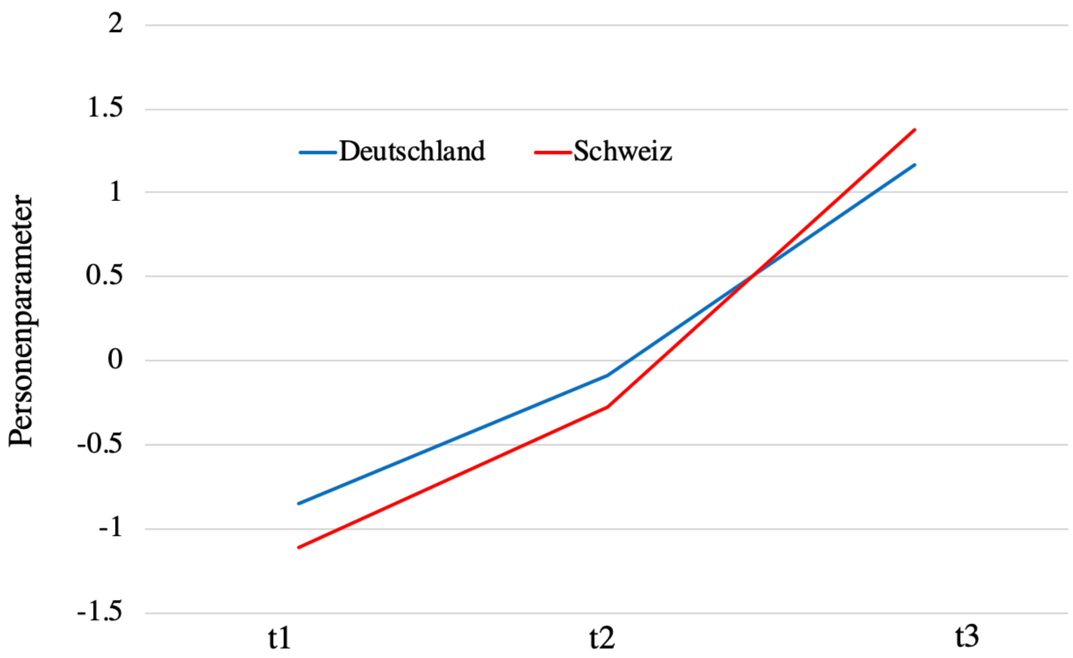

Abbildung 14: Kompetenzentwicklung der Kinder getrennt nach Land

Insgesamt starteten die Kinder in Deutschland mit höheren mathematischen Kompetenzen. Dieser Kompetenzvorsprung blieb auch beim zweiten Testzeitpunkt erhalten. Zwischen dem zweiten und dritten Testzeitpunkt machten jedoch die Kinder in der Schweiz grössere Fortschritte und wiesen beim dritten Testzeitpunkt signifikant höhere mathematische Kompetenzen auf als die Kinder in Deutschland.

8.5 Unterschiede im Entwicklungsverlauf von Kindern mit niedrigem, mittlerem und hohem Vorwissen

Die vierte Forschungsfrage bezieht sich auf den Entwicklungsverlauf von Kindern mit niedrigem, mittlerem und hohem Vorwissen. Zur Gruppe der Kinder mit niedrigem Vorwissen gehörten Kinder, deren mathematische Kompetenzen beim ersten Testzeitpunkt eine Standardabweichung unter dem Mittelwert aufwiesen. Kinder, deren mathematische Kompetenzen beim ersten Testzeitpunkt eine Standardabweichung über dem Mittelwert besassen, gehörten zur Gruppe der Kinder mit hohem Vorwissen. Die mathematischen Kompetenzen der Kinder mit mittlerem Vorwissen lagen zwischen einer Standardabweichung unter und einer Standardabweichung über dem Mittelwert. Tabelle 19 gibt einen Überblick über die Mittelwerte und die Standardabweichungen von Kindern mit unterschiedlichem Vorwissen zu den drei Testzeitpunkten.

Tabelle 19: Übersicht über Mittelwerte und Standardabweichungen von Leistungsgruppen

Vorwissen	T1: *M* (*SD*)	T2: *M* (*SD*)	T3: *M* (*SD*)
niedrig	-3.15 (.58)	-1.98 (1.02)	-.33 (1.22)
mittel	-1.12 (.75)	-.34 (1.24)	1.14 (1.32)
hoch	1.42 (.91)	2.02 (1.34)	3.41 (1.56)

Die Mittelwerte der einzelnen Gruppen unterschieden sich beträchtlich und die Unterschiede waren zu allen drei Testzeitpunkten auch unter Kontrolle von kognitiven Fähigkeiten, Alter, Geschlecht, Erstsprache und Land signifikant: T1 ($F[7,811] = 375.856$, $p = .000$), T2 ($F[7,811] = 145.085$, $p = .000$), T3 ($F[7,811] = 113.504$, $p = .000$). Die Streuung wurde zudem in allen Gruppen grösser. Bei allen drei Testzeitpunkten war die Streuung in der Gruppe mit hohen Leistungen am grössten und in der Gruppe der Kinder mit niedrigen Leistungen am kleinsten.

Zur Überprüfung, ob sich die Leistungsgruppen in ihrer mathematischen Entwicklung über die drei Testzeitpunkte unterscheiden, wurden latente Wachstumskurvenmodelle mit verschiedenen Prädiktoren gerechnet. Dadurch konnte erstens untersucht werden, ob sich die Gruppen bezüglich der Einflussfaktoren auf die Ausgangsleistung bzw. auf den Kompetenzzuwachs über die drei Testzeitpunkte unterschieden. Zweitens konnten die Gruppen bezüglich ihres Ausgangswertes und der Leistungssteigerung verglichen werden. Das Modell hatte akzeptable bis gute Fit-Werte: $CFI = 0.96$; $RMSEA = 0.09$; $SRMR = 0.03$ (vgl. Abschnitt 7.5.5). Tabelle 20 zeigt die Ergebnisse.

Tabelle 20: Ergebnisse des Wachstumskurvenmodelles getrennt nach Leistungsgruppen

	Niedriges Vorwissen				Mittleres Vorwissen				Hohes Vorwissen			
	Intercept		Slope		Intercept		Slope		Intercept		Slope	
Prädiktoren	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Kogn. Fähigk.	.03*	.01	.05***	.01	.05***	.01	.02***	.00	.07***	.02	.01	.01
Geschlecht	-.03	.10	-.15	.09	-.11*	.06	-.07*	.03	-.48**	.15	-.07	.06
Alter	.00	.15	-.20	.13	.23**	.08	-.06	.04	-.07	.23	-.21*	.10
Erstsprache	-.08	.10	-.12	.09	-.36***	.07	-.05	.04	-.37	.21	-.00	.09
Land	.23*	.11	-.35**	.10	.02	.06	-.14***	.03	-.28	.15	-.14*	.07
Ausbildung			-.06	.09			.01	.03			.07	.06
Mittelwert	-3.40***	.80	2.11**	.68	-2.66***	.43	1.04***	.21	1.14	1.25	1.57**	.56
Varianz	.29***	.07	.23***	.05	.45***	.04	.16***	.02	.70***	.12	.22***	.05
Slope mit intercept	*B* = -.09*		*SE* = .04		*B* =- .00		*SE* = .01		*B* = -.03		*SE* = .03	
CFI/RMSEA/SRMR	0.96 / 0.09 / 0.03											
	* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$											

Der Vergleich der drei Gruppen bezüglich beeinflussender Prädiktoren brachte hervor, dass in allen Gruppen die Kinder in der Schweiz einen grösseren mathematischen Kompetenzzuwachs erzielten und dass die kognitiven Fähigkeiten bei allen Kindergruppen einen signifikanten Einfluss auf die Ausgangsleistung hatten. Bei Kindern mit niedrigem und mittlerem Vorwissen beeinflussten die kognitiven Fähigkeiten auch den Kompetenzzuwachs positiv, bei Kindern mit hohem Vorwissen hatten die kognitiven Fähigkeiten auf den Kompetenzzuwachs keinen Einfluss. Nur in der Gruppe der Kinder mit mittlerem Vorwissen wiesen Kinder mit einer anderen Erstsprache als Deutsch beim ersten Testzeitpunkt niedrigere mathematische Kompetenzen auf. In den beiden anderen Gruppen hatte die Erstsprache keinen Einfluss, weder auf die Ausgangsleistung noch auf den Kompetenzzuwachs.

Der Vergleich der drei Gruppen bezüglich des durchschnittlichen Slope-Faktors zeigte, dass die Kinder mit niedrigem Vorwissen den höchsten Anstieg hatten ($B = 2.11$, $p < 0.01$), gefolgt von den Kindern mit hohem Vorwissen ($B = 1.57$, $p < 0.01$). Die Kinder mit mittlerem Vorwissen hatten im Vergleich zu den anderen Gruppen den kleinsten Slope-Faktor ($B = 1.04$, $p < 0.001$).

Weiter wurde ersichtlich, dass die Korrelation von Ausgangsleistung und Leistungssteigerung in der Gruppe der Kinder mit niedrigem Vorwissen signifikant war. Innerhalb dieser Gruppe hatten die Kinder mit niedrigem Vorwissen den grösseren Kompetenzzuwachs.

8.6 Zusammenfassung

Tabelle 21 fasst den Einfluss der Kontextfaktoren auf die mathematische Kompetenz zum ersten Testzeitpunkt und auf die Kompetenzentwicklung zwischen dem ersten und zweiten sowie zwischen dem zweiten und dritten Testzeitpunkt übersichtlich zusammen.

Tabelle 21: Übersicht Einflussfaktoren zu allen drei Testzeitpunkten

	T1			***T2***			***T3***		
	B	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***
Intercept	**-0.79**	0.08	.000	-0.01	0.08	.939	**1.48**	0.09	.000
Individualebene									
MATH^{WLE_t1}				**0.81**	0.03	.000			
MATH^{WLE_t2}							**0.81**	0.02	.000
Alter T1	**0.62**	0.12	.000	-0.04	0.09	.640	-0.16	0.09	.077
Geschlecht	**-0.31**	0.08	.000	**-0.23**	0.06	.000	**-0.14**	0.06	.027
Kogn. Fähigkeiten	**0.15**	0.01	.000	**0.06**	0.01	.000	**0.03**	0.01	.000
Erstsprache	**-0.70**	0.10	.000	**-0.26**	0.08	.002	-0.07	0.08	.397
Klassenebene									
Land	**0.23**	0.10	.016	0.10	0.18	.567	**-0.48**	0.18	.008
Ausbildung				-0.53	0.10	.598	-0.07	0.10	.483
Land*Ausbildung				-0.18	0.20	.367	-0.12	0.20	.557
	Var	***SD***		***Var***	***SD***		***Var***	***SD***	
Zufällige Effekte									
Fehlervarianz L1	1.35	1.16		0.78	0.88		0.77	0.88	
Fehlervarianz L2	0.05	0.22		0.07	0.26		0.07	0.27	
Erklärte Gesamtvarianz (%)	34			69			71		

Zusammenfassend kann festgehalten werden, dass das mathematische Vorwissen die grösste Erklärungskraft für die Unterschiede in den Leistungen zum zweiten und zum dritten Testzeitpunkt besass. Beim ersten Testzeitpunkt leisteten die kognitiven Fähigkeiten den grössten Beitrag zur Erklärung der Kompetenzunterschiede. Die kognitiven Fähigkeiten waren zudem auch für Unterschiede in der Leistungsentwicklung bedeutsam. Je höher die kognitiven Fähigkeiten waren, desto höher waren die mathematischen Leistungen beim ersten Testzeitpunkt und desto grössere Fortschritte haben die Kinder gemacht. Auf der Individualebene erwies sich über alle drei Testzeitpunkte hinweg das Geschlecht als ein signifikanter Prädiktor. Die Jungen starteten mit höheren mathematischen Kompetenzen und erzielten einen grösseren Kompetenzzuwachs. Das Alter hatte nur einen Einfluss auf die Leistung beim ersten Testzeitpunkt. Ältere Kinder hatten dabei höhere Kompetenzen als jüngere. Kinder mit Deutsch als Erstsprache starteten mit höheren Kompetenzen und zeigten zwischen dem ersten und zweiten Testzeitpunkt auch eine positivere Entwicklung als Kinder mit einer anderen Erstsprache als Deutsch. Zwischen dem zweiten und dritten Testzeitpunkt machten Kinder unabhängig ihrer Erstsprache die gleich grossen Fortschritte. Auf der Klassenebene zeigte sich, dass die Kinder in Deutschland beim ersten Testzeitpunkt mit höheren mathematischen Kompetenzen starteten, die Kinder in der Schweiz aber einen grösseren Leistungszuwachs zwischen dem zweiten und dritten Testzeitpunkt hatten. Tendenziell erzielten Kinder, die beim ersten Testzeitpunkt niedrige mathematische Kompetenzen aufweisen einen grösseren Kompetenzzuwachs.

Der Vergleich von Kindern mit niedrigem, mittlerem und hohem Vorwissen bezüglich ihrer mathematischen Leistungsentwicklung liess erkennen, dass die Kinder mit niedrigem Vorwissen den höchsten Slope-Faktor hatten, gefolgt von den Kindern mit hohem Vorwissen. Auch hier hatte das Land in allen drei Gruppen einen signifikanten Effekt. Die Kinder in der Schweiz hatten in allen Gruppen einen grösseren Kompetenzzuwachs als die Kinder in Deutschland.

9 Zusammenfassung und Diskussion

Die vorliegende Studie fokussierte die mathematische Entwicklung von Kindergartenkindern. Dabei stand die Frage nach Unterschieden in der mathematischen Entwicklung im Vordergrund. Im Rahmen einer Längsschnittuntersuchung wurde über drei Messzeitpunkte hinweg geprüft, wie sich die mathematischen Kompetenzen unter verschiedenen Bedingungen entwickeln. Von Interesse waren der Einfluss der Kontextfaktoren kognitive Fähigkeiten, Alter, Erstsprache und Geschlecht auf die Leistungsentwicklung und die Frage, ob Unterschiede in der mathematischen Entwicklung zwischen Kindern in Deutschland und Kindern in der Schweiz bestehen. Ausserdem wurde untersucht, wie sich Kinder mit niedrigem, mittlerem und hohem Vorwissen bezüglich ihrer mathematischen Kompetenzen entwickeln. Zusätzlich wurde im Rahmen dieser Untersuchung überprüft, ob sich ein bestehendes, adaptiertes mathematisches Testinstrument, der TEDI-MATH (Kaufmann et al., 2009), zur längsschnittlichen Kompetenzerfassung von Kindergartenkindern eignet. Dabei interessierten insbesondere die Itemhomogenität und die Messinvarianz über die Zeit und von Subgruppen. Im Folgenden werden zuerst diese Ergebnisse zusammengefasst und diskutiert, bevor dann in den Abschnitten 9.2 und 9.3 die Hypothesen zum mathematischen Entwicklungsverlauf geprüft und diskutiert werden. In Abschnitt 9.4 wird die Untersuchung kritisch reflektiert und deren Limitationen aufgezeigt. Die Arbeit endet mit einem Fazit, welches die wichtigsten Erkenntnisse der Untersuchung festhält.

9.1 Diskussion der Überprüfung des Mathematiktests mit dem Rasch-Modell

Mit der ersten Fragestellung wurde die Eignung des Testinstrumentes zur Messung mathematischer Kompetenzen im Längsschnitt untersucht. Nebst der Überprüfung der Reliabilität (Cronbachs Alpha) wurde der Test auch mit dem Rasch-Modell aus der Item-Response-Theorie überprüft. Dabei wurden die Itemhomogenität sowie die Messinvarianz über die Zeit und über ausgewählte Subgruppen analysiert. Die Ergebnisse bestätigen insgesamt, dass sich die in dieser Untersuchung verwendete adaptierte Form des TEDI-MATH von Kaufmann et al. (2009) zur längsschnittlichen Erfassung grundsätzlich eignet. Trotzdem mussten einige Items aufgrund ungünstiger Infit-MNSQ-Werte oder aufgrund von Messvarianz ausgeschlossen werden. Der Ausschluss dieser Items wird im Folgenden diskutiert und ein Fazit gezogen.

9.1.1 Itemhomogenität

Aufgrund ungünstiger Infit-MNSQ-Werte mussten vier Items ausgeschlossen werden. Das waren zwei Items zum Grössenvergleich und ein Item zum approximativen Mengenvergleich bzw. zur numerischen Inklusion. Bei den Items zum Grössenvergleich handelte es sich um zwei sehr schwierige Items, die insgesamt nur von wenigen Kindern gelöst wurden. Ob die Modellpassung deshalb ungenügend war, kann nicht abschliessend beantwortet werden. Denn andere schwierige Items wiesen gute Fit-Werte auf. Möglich wäre daher, dass die Ratewahrscheinlichkeit den Fit-Wert dieser Items beeinflusst hat. Die Ratewahrscheinlichkeit wurde in dieser Untersuchung gegenüber der Originalversion des TEDI-MATH zwar von 50 auf 25 % verringert (vgl. Abschnitt 7.3.1). Dennoch kann es sein, dass auch Kinder mit niedriger Personenfähigkeit diese schwierigen Items zum Teil richtig gelöst haben, indem sie zweimal richtig geraten haben. Die ebenfalls wegen ungünstiger Fit-Werte ausgeschlossen Items zum approximativen Mengenvergleich und zur numerischen Inklusion hatten eine Ratewahrscheinlichkeit von 50 %. Es ist deshalb anzunehmen, dass auch hier die Itemantwort zu wenig von der Eigenschaftsausprägung der Kinder abhing.

9.1.2 Messvarianz

Zeitbezogene Messvarianz

Von zeitbezogener Messvarianz waren vor allem Items zum Zählen, zum Abzählen, zum Zahlen lesen und zu einfachen Additionsaufgaben betroffen. Die Gründe dafür könnten curricularer Natur sein. Es handelt sich um Kompetenzen, denen in den Bildungs- und Lehrplänen im Kindergarten eine hohe Bedeutung zugeschrieben wird (vgl. Abschnitt 2.1.2). So wird im Rahmenplan der Jugendministerkonferenz in Deutschland der Umgang mit Zahlen und Mengen als bedeutend hervorgehoben (Jugendministerkonferenz, 2004) und auch in den beiden Bildungsplänen von Schleswig-Holstein sowie von Niedersachsen wird die Beschäftigung mit Zahlen und Grössen als wichtige mathematische Tätigkeit in Kindertagesstätten erwähnt (Knauer & Hansen, 2012; Niedersächsisches Kultusministerium, 2011). Auch im Lehrplan der Schweiz sind der Aufbau von Zahl-, Zähl- und Abzähl-Kompetenzen sowie Kompetenzen zum Vergleichen von Anzahlen im Kindergarten verbindlich vorgeschrieben (D-EDK, 2014). Zudem handelt es sich um prozedurale Fähigkeiten, die geübt werden können. Es ist somit möglich, dass diese mathematischen

Tätigkeiten im Vergleich zu anderen gezielter und intensiver gefördert wurden und die Items deshalb von zeitbezogener Messvarianz betroffen waren. Dies lässt sich exemplarisch an zwei Items verdeutlichen. Die Aufgabe zum Zählen in Zweierschritten befand sich beim ersten Testzeitpunkt auf Position 74 innerhalb der aufsteigenden Schwierigkeit aller Items, beim dritten Testzeitpunkt aber nur noch auf Position 64. Das Zählen in Zweierschritten wird während der Kindergartenzeit geübt und ist für viele Kinder letztlich wie das Aufsagen eines Verses. Die Additionsaufgabe 2+2 lag beim ersten Testzeitpunkt noch auf Position 45, beim dritten Testzeitpunkt hingegen nur noch auf Position 27. Aus Untersuchungen ist bekannt, dass Kinder solche einfachen Verdoppelungsaufgaben schnell auswendig wissen und diese für sie viel einfacher zu lösen sind als andere Additionsaufgaben (Gaidoschik, 2010). Es ist deshalb anzunehmen, dass beide Items zum dritten Testzeitpunkt in Relation zu den anderen Items einfacher zu lösen waren als noch beim ersten Testzeitpunkt.

Messvarianz in Bezug auf die Länder

Zu den Items, in denen sich Messvarianz bezüglich des Landes zeigte, gehörten wenige einzelne Items aus verschiedenen Subtests, die oft auch von zeitlicher oder geschlechterspezifischer Messvarianz betroffen waren. Betrachtet man die Items, die nur von DIF (Differential Item Functioning) Land betroffen waren, so waren dies ein Item zum Zählen von 3 bis 10, die beiden Items zum Weiterzählen um eine Anzahl Schritte und alle Items zum Ordnen von Anzahlen und Zahlen nach deren Grösse. Die Messvarianz dieser Items betraf mehrheitlich den dritten, teilweise auch den zweiten und dritten Messzeitpunkt, nicht aber den ersten Testzeitpunkt. Das deutet darauf hin, dass diese mathematischen Tätigkeiten in den beiden Ländern unterschiedlich stark gefördert worden sind. Vergleicht man bei den betroffenen Items wiederum die Position innerhalb der aufsteigenden Itemschwierigkeit wird deutlich, dass die beiden Items zum Weiterzählen um eine bestimmte Anzahl Schritte in Deutschland beim dritten Testzeitpunkt relativ zu den anderen Items einfacher zu lösen waren als in der Schweiz. In Deutschland nahmen diese beiden Items die Positionen 61 bzw. 68 ein, in der Schweiz die Positionen 70 bzw. 76. Anders verhielt es sich bei den Items zum Ordnen von Anzahlen und Zahlen nach der Grösse. Diese Items waren für die Kinder in der Schweiz einfacher zu lösen als für die Kinder in Deutschland. Beispielsweise nahm das Item zum Ordnen von Grössen beim dritten Testzeitpunkt in der Schweiz Position 29 ein und in Deutschland Position 42.

Insgesamt belegen diese Befunde, dass der Kontext, in dem Daten erhoben werden, eine Rolle spielen kann und bei der Konstruktion von Leistungstests allenfalls mitberücksichtigt werden muss. Normen können somit nicht ohne Weiteres von einem Land auf ein anderes übertragen werden.

Messvarianz in Bezug auf die Geschlechter

Vier Items zum Lesen von grossen Zahlen und das Item zur Zahlen-Grössen-Zuordnung wurden ausschliesslich aufgrund von Messvarianz zwischen den Geschlechtern aus den Analysen ausgeschlossen, wobei das Item zur Zahlen-Grössen-Zuordnung nur beim ersten Testzeitpunkt von Subgruppenvarianz betroffen war. Das Item war zu diesem Zeitpunkt für die Mädchen relativ zu den anderen Items einfacher zu lösen als für die Jungen. Die Items zum Lesen von grossen Zahlen waren hingegen durchwegs für die Jungen einfacher zu lösen. Beispielsweise wurden die Zahlen 105 und 160 beim dritten Testzeitpunkt viermal häufiger von Jungen richtig gelesen als von Mädchen. Eine Interpretation dieser Einzelergebnisse gestaltet sich schwierig, insbesondere auch deshalb, weil die Forschungsergebnisse zu Geschlechterdifferenzen beim mathematischen Lernen von jungen Kindern uneinheitlich sind. Während in einigen Studien das Geschlecht keinen Einfluss auf die mathematischen Kompetenzen im Kindergarten hatte (z. B. Dornheim, 2008; Niklas & Schneider, 2012), wurden in anderen signifikante Effekte beobachtet (z. B. Anders et al., 2012; Krajewski, 2008), (vgl. auch Abschnitt 5.3). In der Studie von Krajewski (2008) zeigte sich, dass die Unterschiede zwischen Jungen und Mädchen insbesondere im Zahlwissen und im Zählen vorhanden waren, wobei die Jungen höhere Ergebnisse erzielten. Die hier von geschlechterspezifischer DIF betroffenen Items zum Zahlen lesen lassen sich dem Bereich Zahlwissen von Krajewski zuordnen.

9.1.3 Fazit zum Mathematiktest

In vorliegender Untersuchung wurde der TEDI-MATH von Kaufmann et al. (2009) als mathematisches Testinstrument eingesetzt, wobei einige Anpassungen, insbesondere zur Reduktion der Ratewahrscheinlichkeit bzw. zur Vermeidung eines Deckeneffekts vorgenommen wurden (vgl. Abschnitt 6.3.4). Zusätzlich wurden zwei neue Subtests in den Test aufgenommen, die im TEDI-MATH nicht vorhanden sind, aber der Überprüfung zweier weiterer Kernkompetenzen in der mathematischen Entwicklung dienen. Es waren dies

Aufgaben zur Zuordnung von Zahlen zu Mengen und zur Überprüfung des Verständnisses von Relationalzahlen. Diese beiden Kompetenzen stellen laut den Entwicklungsmodellen von Krajewski (2008) sowie Fritz und Ricken (2008) Meilensteine in der mathematischen Entwicklung dar (vgl. Abschnitt 4.2). Nach Fischer et al. (2017) fehlen im TEDI-MATH zudem Aufgaben zur Kernkompetenz Zahlenraumvorstellung, beispielsweise das Verorten von Zahlen auf einem Zahlenstrahl. Da dieses ordinale Zahlverständnis aber auch in der Aufgabe zum Ordnen von Zahlen nach der Grösse überprüft wird, wurden für diesen Bereich nicht noch zusätzliche Aufgaben in den Test aufgenommen.

Die Analysen mit dem Rasch-Modell geben insgesamt Aufschluss darüber, dass mit der adaptierten Version des TEDI-MATH mathematische Kompetenzen von Kindergartenkindern reliabel erfasst werden können. Die WLE-Reliabilität war für alle Messzeitpunkte sehr hoch und deutlich höher als Cronbachs Alpha in der, allerdings deutlich kleineren, deutschen Normierungsstichprobe. Die Ergebnisse weisen insgesamt darauf hin, dass der angepasste TEDI-MATH zur längsschnittlichen Erfassung mathematischer Kompetenzen von Kindergartenkindern geeignet ist. Obwohl eine grosse Anzahl von Items ausgeschlossen wurde, blieben genügend Items, die es ermöglichten, die mathematischen Leistungen einer sehr heterogenen Gruppe reliabel zu messen und Aussagen über den Lernzuwachs zu treffen.

Trotzdem bleibt zu überlegen, ob durch den Ausschluss von Items zentrale Inhalte mathematischer Kompetenzen im Kindergartenalter, wie beispielsweise das Ordnen von Anzahlen nach der Grösse, nicht mehr im Test vorhanden sind und somit nicht überprüft werden können. In diesem Fall muss eine Entscheidung getroffen werden, ob ein statistisches oder ein inhaltliches Kriterium stärker gewichtet werden soll. Wird der Test beispielsweise zur Diagnostik eingesetzt, z. B. zur Früherkennung von Rechenstörungen, kann es allenfalls von Bedeutung sein, dass Items zu allen mathematischen Kernkompetenzen vorhanden sind. In der vorliegenden Untersuchung war es aber vor allem wichtig, dass der Test die mathematischen Kompetenzen der Kinder insgesamt reliabel erfasst, damit Aussagen über deren mathematische Entwicklung gemacht werden können. Der Ausschluss einzelner Aufgabentypen ist deshalb als unproblematisch zu bezeichnen. Zudem bestätigt die hohe Korrelation der WLEs mit und ohne Ausschluss von unbefriedigenden Items zu allen drei Testzeitpunkten, dass dasselbe Konstrukt gemessen wurde. Dieses Ergebnis kann in Bezug auf die Testökonomie bedeutsam sein, kann doch mit deutlich weniger Items eine zuverlässige längsschnittliche Messung erfolgen.

9.2 Hypothesenprüfung zu den mathematischen Kompetenzen und ihren Einflussfaktoren

Im Zentrum dieser Arbeit standen Fragen nach den mathematischen Kompetenzen von Kindergartenkindern zu einem ersten Testzeitpunkt und deren Entwicklung über drei Testzeitpunkte. Es interessierte, über welche mathematischen Kompetenzen ca. fünfjährige Kinder verfügen, wie stark sie sich in ihren Kompetenzen und ihrer Leistungsentwicklung unterscheiden und welche Faktoren Unterschiede zu erklären vermögen.

9.2.1 Mathematische Leistungen von Kindergartenkindern zum ersten Testzeitpunkt

Die Ergebnisse haben gezeigt, dass die Kinder erwartungsgemäss sehr heterogene mathematische Kompetenzen beim ersten Testzeitpunkt aufwiesen. Die Streuung der Personenparameter reichte von -6.49 bis 5.98 mit einem Mittelwert von -1.06 und einer Standardabweichung von 1.48. Diese grosse Heterogenität wurde in diversen Studien bestätigt (z. B. Hasemann & Gasteiger, 2014; Moser, Stamm & Hollenweger, 2004). Die zehn leichtesten Items wurden von 63.5 % aller Kinder gelöst. Diese Aufgaben sind in den Entwicklungsmodellen von Krajewski (2008) sowie von Fritz und Ricken (2008) den ersten beiden Ebenen bzw. den ersten drei Stufen zuzuordnen. Entsprechend handelt es sich um eher einfache Aufgaben zum Zählen und Abzählen. Trotzdem war eine beachtliche Anzahl von Kindern nicht in der Lage, diese Aufgaben zu lösen. Daneben konnten fast 10 % der untersuchten Kinder Items bearbeiten, die der dritten Ebene im Modell von Krajewski (2008) bzw. der vierten und fünften Stufe im Modell von Fritz und Ricken (2008) zugeordnet sind. Es handelt sich um Items zum Zerlegen von Zahlen, zu Zahlbeziehungen oder zu Additionsaufgaben. Das Lösen der meisten dieser Aufgaben erfordert Kompetenzen, die in den Lehrplänen erst am Ende der ersten Klasse der Grundschule erwartet werden (Buddenberg et al., 2017; D-EDK, 2014). Das heisst, dass fast 10 % der Kinder rund anderthalb Jahre vor dem Schuleintritt bereits über einen beträchtlichen Teil mathematischer Kompetenzen verfügen, die während der ersten Klasse vermittelt werden.

9.2.2 Einfluss des Vorwissens

Aus verschiedenen Studien ist bekannt, dass hohe frühe mathematische Kompetenzen mit späteren hohen mathematischen Leistungen einhergehen (z. B. Hornung et al., 2014; von Aster et al., 2007; Watts et al., 2014). In dieser Studie wurde der Einfluss des mathematischen Vorwissens auf die mathematischen Kompetenzen auch für den Kindergarten bestätigt. Je höher die Kompetenzen beim ersten bzw. zweiten Testzeitpunkt waren, desto höher waren diese auch beim zweiten bzw. dritten Testzeitpunkt. Das Vorwissen erklärte jeweils 55 bzw. 60 % der Unterschiede in den Kompetenzen zum zweiten und dritten Testzeitpunkt. Der Einfluss des Vorwissens auf den Leistungszuwachs wurde knapp nicht signifikant (vgl. Abschnitt 8.3.1). In der Tendenz zeigen aber Kinder mit niedrigen mathematischen Leistungen zum ersten Testzeitpunkt über die gesamte Kindergartenzeit einen grösseren Kompetenzzuwachs (vgl. Abschnitt 9.3). Dennoch indizieren die hohen Korrelationen ($.74 < r < .82$) zwischen den Messzeitpunkten stabile relative Positionen inerhalb der Kindergruppe. Wer zum ersten Testzeitpunkt höhere mathematische Kompetenzen aufwies, gehört auch beim dritten Testzeitpunkt, kurz vor Schuleintritt, mit einer hohen Wahrscheinlichkeit zu den mathematisch leistungsstärkeren Kindern. Damit wurde einmal mehr die Bedeutung des mathematischen Vorwissens für die Vorhersage von späteren mathematischen Kompetenzen belegt. Gleichzeitig weisen die Ergebnisse dieser Untersuchung darauf hin, dass nicht nur die mathematischen Kompetenzen im Kindergarten entscheidend für spätere mathematische Leistungen sind (vgl. Kapitel 3), sondern auch, ob und in welchem Ausmass Kinder bereits vor dem Eintritt in den Kindergarten mathematische Kompetenzen erwerben konnten.

9.2.3 Einfluss der kognitiven Fähigkeiten

Da beim ersten Testzeitpunkt das mathematische Vorwissen noch nicht ins Modell aufgenommen werden konnte, beeinflussten erwartungsgemäss die allgemeinen kognitiven Fähigkeiten der Kinder ihre mathematischen Leistungen. Je höher diese waren, desto höher waren auch die mathematischen Kompetenzen. Mit einer Varianzaufklärung von 24 % stellten die kognitiven Fähigkeiten beim ersten Testzeitpunkt den bedeutsamsten Prädiktor dar. Dieser Befund deckt sich beispielsweise mit der Studie von Sale et al. 2018 oder von Krajewski und Schneider (2009), die den Einfluss der Intelligenz auf die mathematischen Kompetenzen im Kindergarten nachwiesen. Letztere berichtete auch, dass der Einfluss der Intelligenz kleiner wurde, wenn das Vorwissen als Prädiktor mitaufgenommen wurde. Auch in dieser Studie nahm der Einfluss der kognitiven Fähigkeiten unter Einbezug des Vorwissens ab, blieb aber hypothesenkonform signifikant. Die durch die kognitiven Fähigkeiten erklärte Varianz im Kompetenzzuwachs betrug zwischen dem ersten und zweiten Testzeitpunkt 7 % und zwischen dem zweiten und dritten noch 4 %. Dieser Befund stimmt beispielsweise mit der vom Design her vergleichbaren Längsschnittstudie von Hauser et al. (2014) überein. Ein ähnliches Design besass auch die Studie von Schuchardt et al. (2014). Hier leistete die Intelligenz keinen Beitrag zur Erklärung von Unterschieden in der Kompetenzentwicklung von Kindergartenkindern. Allerdings erwiesen sich das visuell-räumliche Arbeitsgedächtnis und die Benenngeschwindigkeit als signifikante Prädiktoren. Es ist anzunehmen, dass diese hoch mit der Intelligenz korrelieren, handelt es sich doch dabei auch um kognitive Fähigkeiten. Möglicherweise wurde die Variable Intelligenz aufgrund von Multikollinearität nicht signifikant.

Zusammenfassend lässt sich festhalten, dass die kognitiven Fähigkeiten durchaus einen Einfluss auf die mathematischen Kompetenzen und deren Entwicklung haben, dass aber der Einfluss des Vorwissens bedeutend grösser ist.

9.2.4 Einfluss des Geschlechts

Im Vorfeld wurde erwartet, dass das Geschlecht keinen Einfluss auf die mathematischen Leistungen zum ersten Testzeitpunkt hat. Diese These muss jedoch verworfen werden. Entgegen verschiedener Studien, die einen Einfluss des Geschlechts auf die mittleren mathematischen Leistungen erst in der Grundschule nachwiesen (Dornheim, 2008; Niklas & Schneider, 2012; Sahr, 2012), zeigten die Jungen in dieser Untersuchung bereits beim ersten Testzeitpunkt im Kindergarten höhere mathematische Leistungen. Obwohl verschiedenste Studien keine geschlechterspezifischen Unterschiede in den frühen mathematischen Kompetenzen feststellen konnten, gibt es auch Hinweise darauf, dass die Jungen bereits im Kindergarten gegenüber den Mädchen höhere Kompetenzen besitzen. Bei der Testnormierung beim MBK O (Krajewski, 2018) beispielsweise liessen die Jungen teilweise signifikante Vorteile gegenüber den Mädchen erkennen (vgl. Abschnitt 6.3.5). Auch in der Studie von Krajewski (2009) zeigten die Jungen im zweiten Kindergartenjahr rund sechs Monate vor Schuleintritt höhere Leistungen im Bereich Zahlenvorwissen. Die Unterschiede liessen sich vor allem auf das arabische Zahlenwissen und das Zählen zurückführen. Auch in dieser Untersuchung wiesen die Jungen bei drei Aufgaben zum Zahlen lesen

($\chi^2_{(df=1)}$= 14.09, p = .000; $\chi^2_{(df=1)}$= 9.47, p = .002; $\chi^2_{(df=1)}$= 14.68, p = .000, n = 894) und bei der Aufgabe zum rückwärts Zählen ($\chi^2_{(df=1)}$= 5.20, p = .023, n = 894) einen signifikanten Vorsprung auf, aber auch bei Additionsaufgaben ($\chi^2_{(df=1)}$= 10.37 p = .001, n = 894), bei Aufgaben zu Zahlbeziehungen ($\chi^2_{(df=1)}$= 6.72, p = .010; $\chi^2_{(df=1)}$= 6.73, p = .009; $\chi^2_{(df=1)}$= 5.09, p = .024, n = 894) oder bei einer Aufgabe zum Grössenvergleich von Zahlen ($\chi^2_{(df=1)}$= 4.67, p = .031, n = 894). Es handelt sich mehrheitlich um Aufgaben, die in den Modellen der Zahlbegriffsentwicklung auf den höchsten Ebenen bzw. Stufen anzusiedeln sind (vgl. Abschnitt 4.2). Es scheint also, als hätten die Jungen insgesamt höhere mathematische Kompetenzen beim ersten Testzeitpunkt, die vor allem auf das Lösen von komplexeren Aufgaben zurückzuführen sind. Zudem betreffen die Unterschiede alles Aufgaben, die auf symbolischer Repräsentationsebene angeboten wurden (vgl. Abschnitt 4.3.5). Das heisst, dass Jungen im Vergleich zu Mädchen rein formale Aufgaben häufiger lösen konnten.

Was den Einfluss des Geschlechts auf die mathematische Kompetenzentwicklung im Längsschnitt angeht, zeigte sich hypothesenkonform, dass die Jungen nicht nur mit höheren Kompetenzen starteten, sondern auch einen höheren Leistungszuwachs über die drei Testzeitpunkte erzielten. Dieser Befund deckt sich mit den beiden Längsschnittstudien von Anders et al. (2012) und Weinhold Zulauf et al. (2003). Auch dort zeigten die Jungen einen grösseren mathematischen Kompetenzzuwachs im Kindergarten.

Worauf diese geschlechterspezifischen Unterschiede zurückzuführen sind, lässt sich nicht abschliessend beantworten. Es gibt jedoch Hinweise darauf, dass geschlechterbedingte Unterschiede in der Erziehung, beispielsweise eine höhere Erwartungshaltung von Bezugspersonen bezüglich mathematischer Fähigkeiten (vgl. Abschnitt 5.3), die Unterschiede erklären können (Hyde et al., 2008). Aus Untersuchungen ist auch bekannt, dass Jungen bereits im Kindergarten ein höheres Selbstkonzept in Bezug zu mathematischen Tätigkeiten aufweisen (Jacobs et al., 2002; Marsh et al., 2002) (vgl. Abschnitt 5.3). Der Zusammenhang zwischen höherem Selbstkonzept und höheren Kompetenzen konnte in mehreren Studien auch bereits für die Vorschule nachgewiesen werden, insbesondere dann, wenn Selbstkonzept und Leistung fachspezifisch erhoben wurden (Mantzicopoulos, 2006; Valeski & Stipek, 2001). In Bezug auf Mathematik konnten zwei Studien bei vier- bis siebenjährigen Kindern belegen, dass ein höheres Selbstkonzept mit höheren mathematischen Leistungen einhergeht (Abt Gürber, 2011; Marsh et al., 2002). In Bezug auf längsschnittliche Unterschiede (Kompetenzzuwachs) können allenfalls geschlechterspezifische Stereotype der pädagogischen Fachkräfte zur Erklärung beigezogen werde. So konnten Holder und Kessels (2017) in ihrer Studie bestätigen, dass angehende Lehrpersonen der Grundschule mehrheitlich die Überzeugung teilen, dass Mädchen in Mathematik weniger gut sind als Jungen. Auch Tiedemann (2000) zeigte, dass geschlechterspezifische Unterschiede in den mathematischen Kompetenzen von der Erwartungshaltung bzw. der Einstellung von Lehrpersonen beeinflusst sind. Demnach bewerteten Lehrpersonen das Erreichen von mathematischen Lernzielen für Mädchen im Gegensatz zu Jungen als schwieriger und ungenügende Leistungen bei Mädchen wurden auf niedrige Fähigkeiten zurückgeführt, während bei Jungen eine zu geringe Anstrengung hierfür verantwortlich gemacht wurde. Im Wissen, dass Kinder im Alter von vier bis sechs Jahren auf Leistungserwartungen von Lehrpersonen besonders sensibel reagieren (Babad, 2010; Hattie, Beywl & Zierer, 2013), könnten solche Stereotypen auch eine mögliche Erklärung für geschlechterspezifische Disparitäten in der mathematischen Kompetenzentwicklung von Kindergartenkindern sein. Daneben sind aber auch andere Einflussfaktoren wie beispielsweise Interessensunterschiede zwischen Jungen und Mädchen denkbar. So vermuteten Weinhold Zulauf et al. (2003) in ihrer Untersuchung, dass das unterschiedliche Spieleinteresse bei Jungen und Mädchen eine grosse Bedeutung für die Unterschiede in der mathematischen Kompetenzentwicklung hat. Sie folgerten, dass im stärker wettbewerbsbezogenen Spiel der Jungen Zahlen wichtige Masseinheiten darstellen, wohingegen das stärkere beziehungs- und phantasiebezogene Spiel der Mädchen weniger Übungsmöglichkeiten im Bereich Zahlen bietet. Ob die geschlechterbedingten Unterschiede in den mathematischen Leistungen und der Kompetenzentwicklung auf geschlechterspezifische Erwartungshaltungen von Bezugs- und Erziehungspersonen und/oder auf Unterschiede im Selbstkonzept oder im Interesse von Jungen und Mädchen zurückzuführen sind, lässt sich nicht abschliessend beantworten. Hierzu würden weitere Forschungsarbeiten benötigt.

Unabhängig davon dürfen diese geschlechterbedingten Differenzen aber auch nicht überbewertet werden. Die Varianzaufklärung durch das Geschlecht war in dieser Studie mit 1 bzw. 3 % relativ gering. Zudem sind die Unterschiede in Bezug auf mathematische Kompetenzen innerhalb der Jungen-Gruppe bzw. der Mädchen-Gruppe bedeutender zu werten als die Divergenz zwischen den Geschlechtern. So reichten die Personenparameter der Jungen beim ersten Testzeitpunkt von -5.45 bis 5.53 (M = -.93, SD = 2.48) und die der Mädchen von - 5.45 bis 2.94 (M = -1.12, SD = 1.69). Auch längsschnittlich betrachtet sind die Unterschiede im Kompetenzzuwachs innerhalb der Geschlechter von grösserer Bedeutung als zwischen den

Geschlechtern. Die Leistungszunahme vom ersten zum dritten Testzeitpunt variiert bei den Jungen zwischen - 1.58 und 7.3 Logits (M = 2.37, S = 1.21) und bei den Mädchen zwischen - .95 bis 6.04 (M = 2.21, SD = 1.07). Vor diesem Hintergrund verliert der geschlechterspezifische Unterschied an Bedeutung. Andere Prädiktoren sind indes wichtiger.

9.2.5 Einfluss des Alters

Analog zu anderen Studien (z. B. Jörns et al., 2013 oder Sale et al., 2018) wurde erwartet, dass ältere Kinder beim ersten Testzeitpunkt höhere mathematische Kompetenzen haben als jüngere. Diese Annahme wurde so bestätigt und kann sicherlich auch mit der Streuung des Alters der Stichprobe begründet werden. Diese betrug ca. drei Jahre bei einem Mittelwert von M = 5.20 und einer Standardabweichung von SD = 0.37. Da sich Kinder in diesem Alter in Bezug auf ihre mathematischen Kompetenzen schnell entwickeln (vgl. Abschnitt 4.2), ist nachvollziehbar, dass der grosse Altersunterschied zu Differenzen führen kann. Allerdings wurde auch erwartet, dass ältere Kinder einen höheren Kompetenzzuwachs haben, analog zur Studie von Anders et al. (2012). Dies wurde in der vorliegenden Untersuchung jedoch nicht bestätigt. Das Alter hatte keinen Einfluss auf die mathematische Kompetenzentwicklung, wobei zwischen dem zweiten und dritten Testzeitpunkt tendenziell die jüngeren Kinder grössere Fortschritte machten (vgl. Abschnitt 8.3.3). Denkbar ist dahingehend, dass sich die pädagogischen Fachkräfte speziell der Förderung der jüngeren Kinder angenommen haben, um im Hinblick auf den Übertritt in die Schule deren Leistungsrückstand aufholen zu können.

9.2.6 Einfluss der Erstsprache

Die Ergebnisse der Mehrebenenanalysen haben gezeigt, dass Kinder mit Erstsprache Deutsch beim ersten Testzeitpunkt höhere mathematischen Kompetenzen haben als Kinder mit einer anderen Erstsprache. Dieser Befund wurde so erwartet und deckt sich mit den Ergebnissen der Studie von Sale et al. (2018) oder Anders et al. (2012). Auch hier zeigten sich im Ausgangsniveau Unterschiede, wenn die Familiensprache nicht Deutsch war bzw. beide Elternteile nicht Deutsch als Muttersprache hatten. Diese Ergebnisse sind nachvollziehbar, konnten doch verschiedenste Studien bereits einen positiven Zusammenhang zwischen sprachlichen und mathematischen Kompetenzen nachweisen (Prediger et al., 2015; Rösch & Paetsch, 2011; Ufer & Mehringer, 2013). In Bezug auf den Kindergarten konnten Moser Opitz et al. (2010) belegen, dass Kinder mit einer anderen Erstsprache als Deutsch geringere verbale Zählkompetenzen aufweisen als Kinder mit Erstsprache Deutsch (vgl. Abschnitt 5.5.3). Da diese ein bedeutender Prädiktor für mathematische Kompetenzen sind (Jordan et al., 2007), dürften Kinder mit einer anderen Erstsprache als Deutsch bezüglich deren Entwicklung benachteiligt sein. Auch in dieser Untersuchung waren signifikante Differenzen bezüglich der Zählkompetenzen von Kindern mit Erstsprache Deutsch und Kindern mit einer anderen Erstsprache zu verzeichnen. So konnten zum ersten Testzeitpunkt 39 % aller Kinder mit Erstsprache Deutsch bis 20 und 23 % bis 31 zählen. Bei den Kindern mit einer anderen Erstsprache als Deutsch waren es deutlich weniger. Hier konnten nur 23 % bis 20 bzw. 11 % bis 31 zählen ($\mathcal{X}^2_{(df=2)} = 47.16, p = .000, n = 867$). Beim Rückwärtszählen zeigte sich ein ähnliches Bild. 40 % der Kinder mit Erstsprache Deutsch konnten von 7 bis 1 rückwärts zählen, bei den Kindern mit einer anderen Erstsprache waren es nur 23 % ($\mathcal{X}^2_{(df=1)} = 17.34, p = .000, n = 867$). Da das korrekte Rezitieren der Zahlwortreihe eine wesentliche Voraussetzung bildet, um Anzahlen korrekt zu bestimmen und das Kardinalprinzip zu erwerben (vgl. Abschnitt 4.1.2), dürften die geringeren verbalen Zählkompetenzen ein Erklärungsgrund für die generell niedrigeren mathematischen Kompetenzen sein. Daneben muss berücksichtigt werden, dass Kinder mit einer anderen Erstsprache als Deutsch oft aus Familien mit einem Migrationshintergrund und oft auch mit niedrigem sozioökonomischem Status stammen (Kristen & Granato, 2007). Seit den Anfängen der Bildungs- und Ungleichheitsforschung ist bekannt, dass Familien mit geringem ökonomischem, kulturellem und sozialem Kapital weniger in der Lage sind, ihren Kindern die Kenntnisse und Fertigkeiten zu vermitteln, die sie für eine erfolgreiche Schulkarriere benötigen (Bourdieu, 1983). Da das elterliche Unterstützungsverhalten einen Einfluss auf die frühen mathematischen Kompetenzen hat (vgl. Abschnitt 5.5.4), sind Kinder, deren Erstsprache nicht Deutsch ist, allenfalls auch in dieser Hinsicht benachteiligt.

Im Vorfeld der Studie wurde angenommen, dass die Erstsprache auch einen Einfluss auf die mathematische Leistungsentwicklung hat und Kinder mit Erstsprache Deutsch einen höheren Kompetenzzuwachs aufweisen als Kinder mit einer anderen Erstsprache. Vom ersten zum zweiten Testzeitpunkt hatten Kinder mit Erstsprache Deutsch einen höheren Kompetenzzuwachs als ihre Peers mit einer anderen Erstsprache, nicht aber vom zweiten zum dritten Testzeitpunkt. Im Gegensatz zur Studie von Anders et al. (2012) entwickelten sich die mathematischen Kompetenzen vom zweiten zum dritten Testzeitpunkt unabhängig von der Erstsprache bei allen Kindern gleich. Mit Blick auf die obige Diskussion der Benachteiligung von Kindern mit Migrationshintergrund sind die Ergebnisse dieser Studie erst einmal ermutigend. Zumindest scheint es gelungen zu sein, mit den Regelspielen zur mathematischen Förderung Kinder mit unterschiedlichen Erstsprachen gleichermassen zu fördern. Die Forderung nach mehr Chancengerechtigkeit von sogenannten benachteiligten Kindern aus Familien mit niedrigem sozioökonomischem Status, mit Migrationshintergrund und mit wenig häuslicher Lernanregung durch vorschulische Einrichtung (Becker & Reimer, 2010) wurde allerdings nur teilweise erfüllt. Die Kinder mit einer anderen Erstsprache als Deutsch machten zwar die gleichen mathematischen Fortschritte, der Leistungsrückstand blieb aber erhalten. Der durch die Förderung im Kindergarten diskutierte Ausgleich sozialer Disparitäten (Becker & Reimer, 2010) fand zumindest in Bezug auf die mathematischen Kompetenzen nicht statt.

9.2.7 Einfluss des Landes und der Ausbildung der pädagogischen Fachkräfte

Es wurde erwartet, dass sich die Kinder in Deutschland und in der Schweiz in ihren mathematischen Kompetenzen beim ersten Testzeitpunkt nicht unterscheiden. Diese Hypothese muss mit Blick auf die Ergebnisse der Studie allerdings verworfen werden. Die Kinder in Deutschland hatten im Durchschnitt höhere mathematische Kompetenzen als die Kinder in der Schweiz. Worauf diese Unterschiede zurückzuführen sind, bleibt indessen ungeklärt. Möglich ist, dass sich die Dauer des Kindergartenaufenthaltes positiv auf die mathematischen Kompetenzen ausgewirkt hat. So zeigte sich in der Studie von Anders et al. (2012) ein grosser mathematischer Kompetenzzuwachs der Kinder zwischen drei und fünf Jahren, die den Kindergarten besuchten. Anders wie in Deutschland, wo die Kinder die Kindertagesstätte bereits ab dem ersten Lebensjahr besuchen können, ist der Besuch des Kindergartens in der Schweiz ist erst mit vollendetem viertem Lebensjahr möglich. Das bedeutet, dass die Kinder in der Schweiz beim ersten Testzeitpunkt erst seit ca. einem halben Jahr im Kindergarten waren. Da aber nicht erfasst wurde, wann die Kinder in Deutschland in die KITA eingetreten sind, können dazu keine statistischen Aussagen gemacht werden. Berücksichtigt werden muss jedoch der Umstand, dass die Stichproben nicht repräsentativ für das jeweilige Land sind (vgl. Abschnitt 9.4). Insbesondere in Deutschland wurden nur Klassen aus dem Raum Kiel und Vechta in die Untersuchung einbezogen. Es ist demnach durchaus möglich, dass mit dem Einbezug von Kindergärten aus anderen Regionen oder Bundesländern andere Ergebnisse erzielt worden wären.

Betrachtet man die Entwicklung der mathematischen Kompetenzen im Längsschnitt, zeigte sich entgegen den Erwartungen, dass die Kinder in der Schweiz zwischen dem zweiten und dritten Testzeitpunkt einen signifikant höheren Kompetenzzuwachs hatten. Während dieser Zeit erhielten alle Klassen zehn Regelspiele zur mathematischen Förderung (vgl. Abschnitt 7.2). Die Fachkräfte wurden angeleitet, die Spiele ca. zwei- bis dreimal pro Woche einzusetzen. Es wurde davon ausgegangen, dass dadurch alle Kinder in gleichem Mass eine mathematische Förderung erhalten und sich deshalb keine Unterschiede in der Kompetenzentwicklung der Kinder in Deutschland und der Schweiz zeigen. Diese Hypothese muss jedoch verworfen werden. Der mittlere mathematische Kompetenzzuwachs der Kinder in der Schweiz war signifikant grösser als in Deutschland. Der Anteil an erklärter Varianz auf Klassenebene durch das Land betrug 22 %. Auch die durchschnittliche mathematische Leistung beim dritten Testzeitpunkt war bei den Kindern in der Schweiz unter Kontrolle von Alter, Geschlecht, kognitiven Fähigkeiten und Erstsprache signifikant höher als bei den Kindern in Deutschland. Es ist naheliegend, mögliche Gründe für diese unterschiedliche Kompetenzentwicklung im Einsatz der Spiele im Unterricht zu suchen. Forschungsergebnisse haben hervorgebracht, dass die pädagogischen Fachkräfte in der Schweiz eine passive Lernbegleitung im Spiel im Gegensatz zu den pädagogischen Fachkräften in Deutschland deutlicher ablehnten (Link et al., 2017). Das könnte bedeuten, dass sie aufgrund dieser Überzeugung die erhaltenen Spiele zur mathematischen Förderung eher aktiv begleitet und damit eine grössere Wirkung in Bezug auf den mathematischen Kompetenzzuwachs erreicht haben. Die Bedeutung der Spielbegleitung für das mathematische Lernen haben die Ergebnisse der Studie von Wullschleger (2017) veranschaulicht, in der darauf hingewiesen wird, dass es sinnvoll ist, die Durchführung von Regelspielen zur mathematischen Förderung mit einer individuell-adaptiven Lernunterstützung zu begleiten. Die Studie von Link et al. (2017) hat zudem gezeigt, dass die Lehrpersonen in der Schweiz zielgerichtete und geplante mathematische Lernaktivitäten bevorzugen, während die pädagogischen Fachkräfte in Deutschland eine spontane mathematische Förderung innerhalb des Kindergartenalltags als wichtig erachten. Es könnte also auch sein, dass die pädagogischen Fachkräfte in der Schweiz die Spiele häufiger und konsequenter eingesetzt haben.

Weitere Erklärungsgründe für die divergierende mathematische Kompetenzentwicklung können auch in den länderspezifischen Unterschieden der frühen Bildung liegen (vgl. Abschnitt 2.1). Der Kindergarten in der Schweiz wurde schon sehr früh dem Schulwesen zugewiesen und gehört mittlerweile in fast allen Kantonen zur obligatorischen Schulbildung. Damit verbunden ist ein Bildungsauftrag, der auch die mathematische Förderung miteinbezieht und klare Kompetenzerwartungen zu verschiedenen Bereichen der Mathematik formuliert. Die pädagogischen Fachkräfte im Kindergarten sind dazu verpflichtet, ein entsprechendes Förderangebot bereitzustellen, damit alle Kinder die Minimalziele erreichen. In Deutschland hingegen gehört der Kindergarten nach wie vor zum Sozialbereich und hat vor allem einen Betreuungs- und Erziehungsauftrag. Obwohl unterdessen in praktisch allen Bundesländern Orientierungs- und Bildungspläne für vorschulische Institutionen vorliegen, sind die darin formulierten Bildungsziele im mathematischen Bereich weniger konkret und auch weniger verbindlich formuliert (vgl. Abschnitt 2.1.2). Die Ergebnisse der BiKS-Studie (Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter) belegen zudem, dass in deutschen Kindergärten generell wenig kognitive Förderung stattfindet und insbesondere mathematische Kompetenzen wie die Förderung der Zählfertigkeiten oder der Aufbau des Zahlwissens nur selten vorkommen (Smith, 2013). Es wäre durchaus denkbar, dass die pädagogischen Fachkräfte in der Schweiz aufgrund des verbindlichen Lehrplans den Fokus stärker auf die mathematische Förderung legen als ihre deutschen Kolleginnen und Kollegen. Dies würde allenfalls auch Unterschiede in der Höhe der Intraklassenkorrelation (ICC) zwischen den Ländern erklären. In der Schweiz betrug die erklärte Varianz durch die Klassenzugehörigkeit beim dritten Testzeitpunkt 5 %, in Deutschland war sie mit 8 % deutlich höher (vgl. Abschnitt 8.2.2). Es scheint folglich, als habe die Klassenzugehörigkeit in Deutschland einen grösseren Einfluss auf die mathematischen Leistungen der Kinder als in der Schweiz, allenfalls bedingt durch oben beschriebene Unterschiede in der frühen Bildung. Ob und wie stark mathematische Förderung in den Kindergärten stattfindet, könnte in Deutschland aufgrund weniger verbindlicher Bildungsziele stärker von individuellen Interessen der pädagogischen Fachkräfte oder der spezifischen Ausrichtung der Kindertagesstätte abhängen als in der Schweiz.

Darüber hinaus könnte die unterschiedliche mathematische Kompetenzentwicklung in den beiden Ländern auch auf die Unterschiede in der Ausbildung zur pädagogischen Fachkraft zurückgeführt werden. Es zeigte sich zwar erwartungsgemäss, dass die Ausbildung, gemessen durch die Unterscheidung akademisch versus nicht akademisch, nicht zur Erklärung der Unterschiede in der mathematischen Kompetenzentwicklung auf Klassenebene beitragen konnte. Nebst dem Akademisierungsgrad bestehen aber noch weit bedeutendere Unterschiede in der Ausbildung, die einen Beitrag zur Erklärung leisten könnten. So unterscheiden sich die beiden Ausbildungsinstitutionen *Pädagogische Hochschule* in der Schweiz und *Fachschule für Sozialpädagogik* in Deutschland erheblich voneinander. Wie in Abschnitt 2.1.3 aufgearbeitet,

sind die Ausbildungsthemen und Schwerpunkte der Ausbildung kaum miteinander vergleichbar.
In der Schweiz werden die Kindergärtnerinnen und Kindergärtner zusammen mit den Lehrpersonen der Primarschule ausgebildet. Damit liegt der Fokus stärker auf der Bildungsorientierung, während in Deutschland die sozialpädagogische Ausrichtung im Vordergrund steht. Auch in Bezug auf die Vermittlung mathematischer Themen bestehen erhebliche Unterschiede. In der Schweiz werden spezifische, fachdidaktische Themen in insgesamt vier Modulen über zwei Jahre erarbeitet, während in Deutschland der Lehrplan sehr allgemein für alle Bildungsfächer vorgibt, dass fachdidaktische und sozialpädagogische Kompetenzen zu erwerben sind. Es ist also durchaus denkbar, dass auch die verschiedenen Orientierungen in der Ausbildung zur Folge haben, dass die pädagogischen Fachkräfte in ihren Kindergärten andere Prioritäten in Bezug auf Erziehung oder die Vermittlung von Inhalten setzen.

9.3 Hypothesenprüfung zum Entwicklungsverlauf der Gesamtstichprobe und von Kindern mit unterschiedlichem Vorwissen

Im Folgenden werden zuerst die Hypothesen zum Entwicklungsverlauf der Gesamtstichprobe überprüft. Anschliessend werden die Ergebnisse der Analysen zur Leistungsstreuung und zur Entwicklung von Kindern mit niedrigem, mittlerem und hohem Vorwissen diskutiert, gefolgt von einem Ausblick auf den Übertritt in die Schule.

9.3.1 Entwicklungsverlauf der Gesamtstichprobe

Im Vorfeld wurde gemutmasst, dass die Entwicklung der mathematischen Kompetenzen über die drei Testzeitpunkte analog zu Forschungsarbeiten aus der Grundschule (Ditton & Krüsken, 2009; Karst & Lipowsky, 2013) linear verläuft. Die Ergebnisse der Analyse mit dem Wachstumskurvenmodell liessen jedoch keine lineare Entwicklung erkennen. Die Kinder hatten zwischen dem zweiten und dem dritten Testzeitpunkt einen signifikant grösseren Kompetenzzuwachs als zwischen dem ersten und dem zweiten. Allerdings erhielten die Klassen in dieser Zeit zehn Regelspiele zur mathematischen Förderung. Es ist daher möglich, dass der grössere Kompetenzzuwachs auf den Einsatz dieser mathematischen Spiele zurückzuführen ist.

Die in dieser Studie verwendeten Spiele wurden mit zwei Ausnahmen schon in der Studie von Hauser et al. (2014) eingesetzt und haben sich dort als wirksam erwiesen. Die Kinder, die mit den Spielen gefördert wurden, erzielten im Vergleich zur Kontrollgruppe einen signifikant höheren Lernzuwachs. Auch andere Studien haben die positive Wirkung von Regelspielen zur mathematischen Förderung im Vorschulter nachgewiesen (Jörns et al., 2013; Siegler & Ramani, 2009; Whyte & Bull, 2008). Da in dieser Untersuchung kein Vergleich mit einer Kontrollgruppe möglich war, kann der grössere Kompetenzzuwachs vom zweiten zum dritten Testzeitpunkt jedoch nicht abschliessend auf die Spiele zurückgeführt werden. Denkbar sind auch andere Erklärungen. Beispielsweise könnten die pädagogischen Fachkräfte im Hinblick auf den Übertritt in die Schule den Fokus verstärkt auf die mathematische Förderung gelegt haben.

9.3.2 Leistungsstreuung und Entwicklung von Kindern mit unterschiedlichem Vorwissen

Es wurde auch angenommen, dass die Leistungsstreuung analog zu Studien der Grundschule (Ditton & Krüsken, 2009; Karst & Lipowsky, 2013) über die drei Testzeitpunkte abnimmt. Diese Hypothese muss indes ebenfalls verworfen werden. Die Leistungsstreuung im Kindergarten hat über die drei Testzeitpunkte zugenommen. Die Autoren obiger Studien aus der Grundschule erklären die Reduzierung der Streuung unter anderem mit der egalisierenden Tendenz des Grundschulunterrichts und damit, dass Lehrkräfte die grosse Heterogenität mit einem Leistungsausgleich minimieren wollen, indem sie insbesondere die leistungsschwächeren Schülerinnen und Schüler stärker fördern. Es scheint, als zeige sich diese Tendenz im Kindergarten nicht. Möglich wäre, dass die Fachkräfte im Kindergarten im Gegensatz zu den Lehrpersonen der Grundschule weniger einen Anspruch auf Leistungsausgleich haben. Um diese Vermutung abzusichern, bräuchte es aber zusätzliche Informationen der Fachkräfte, beispielswiese zu ihren Überzeugungen zur mathematischen Förderung im Kindergarten oder zum Umgang mit Heterogenität. Denkbar ist auch, dass der Einsatz von obligatorischen Lehrmitteln bzw. Schulbüchern einen Einfluss auf die Leistungsstreuung hat. Während im Kindergarten keine mathematischen Schulbücher eingesetzt werden, könnten diese in der Grundschule einen Unterricht im Gleichschritt begünstigen und deshalb leistungsausgleichend wirken können. Möglich ist auch, dass die in dieser Studie eingesetzten Spiele die hohe Leistungsstreuung mitbeeinflusst haben. Die Spiele fördern mathematische Kompetenzen auf verschiedenen Niveaus (vgl. Abschnitt 7.2). Auch

Kinder mit hohen mathematischen Kompetenzen erhielten somit mathematische Anregungen. Allenfalls hat sich der sogenannte Matthäus-Effekt (z. B. Zuckermann & Merton, 2010) gezeigt, der besagt, dass Kinder mit höheren Kompetenzen stärker von einer Förderung profitieren und anfangs bestehende Leistungsunterschiede somit nicht kleiner, sondern grösser werden.

Die Tendenz der zunehmend grösseren mathematischen Leistungsstreuung über die zwei Kindergartenjahre verdeutlicht auch, dass sich der Effekt der Regression zur Mitte bei wiederholender Messung (Nachtigall & Suhl, 2002) nur für die Gruppe der Kinder mit niedrigen Kompetenzen, nicht aber für die Gruppe mit hohen Kompetenzen feststellen lässt. Unter Regression zur Mitte wird die Tendenz verstanden, dass sich die Extremausprägungen von Gruppen bei mehrmaliger Messung abschwächen und sich ihre Messungen der mittleren Leistung annähern (Döring & Bortz, 2016). Allerdings ist der grosse mathematische Kompetenzzuwachs der Kinder mit niedrigem Vorwissen in dieser Studie nicht allein mit dem Regressionseffekt erklärbar. Im Gegenteil – „der Regressionseffekt ist ein messfehlerbedingtes statistisches Artefakt und fällt deshalb umso stärker aus, je weniger reliabel die Messungen sind. Mit hoch zuverlässigen Messverfahren kann man also den Regressionseffekt verringern" (Rost, 2013, S. 104). Das in dieser Untersuchung verwendete Messinstrument erfüllt die Kriterien hoher Reliabilität. Zudem weisen auch die hohen Korrelationen der drei Messwerte auf eine, wenn überhaupt, sehr geringe Regression zur Mitte hin.

Der grössere Kompetenzzuwachs der Kinder mit niedrigem Vorwissen kann allenfalls damit erklärt werden, dass die pädagogischen Fachkräfte diese Kinder besonders gefördert haben, um den Leistungsrückstand zu minimieren. Möglich ist darüber hinaus, dass sich die eingesetzten Spiele zur mathematischen Förderung in dieser Gruppe besonders positiv ausgewirkt haben. So zeigte beispielsweise die Untersuchung von Seeger, Holodynski und Roth (2018), dass Kinder mit einem diagnostizierten Entwicklungsrisiko bezüglich mathematischer Kompetenzen im Kindergarten durch den Einsatz einer Sammlung von Brett-, Karten- und Bewegungsspielen mit mathematischem Inhalt gegenüber der Kontrollgruppe signifikante Fortschritte erzielten. Auch in der Studie von Jörns, Schuchardt, Grube und Mähler (2014) konnten Kinder, deren mathematische Leistungen unter dem Median der Gesamtstichprobe lagen, mit zahl- und mengenbezogenen Spielen gefördert werden. Sie wiesen gegenüber der leistungsstärkeren Kontrollgruppe einen grösseren Kompetenzzuwachs auf. Allerdings muss festgehalten werden, dass in dieser Untersuchung trotz grösserem Kompetenzzuwachs der Leistungsrückstand insgesamt erhalten blieb und Kinder mit niedrigen mathematischen Kompetenzen zum ersten Testzeitpunkt auch kurz vor dem Übertritt in die Schule zu den leistungsschwächeren Kindern gehörten (vgl. Abschnitt 9.2.2). Folglich können sie auch im Hinblick auf spätere schulische Mathematikleistungen benachteiligt sein (vgl. Abschnitt 3.2.2).

9.3.4 Ausblick auf den Schulübertritt

In Bezug auf den Übertritt in die Grundschule sind die Ergebnisse zur grossen mathematischen Leistungsstreuung am Ende der Kindergartenzeit von Bedeutung. Es wird hinsichtlich der mathematischen Kompetenzen eine sehr heterogene Gruppe in die erste Klasse eintreten. Fast 10 % der Kinder konnten am Ende des Kindergartens noch nicht bis 20 zählen und 14 % noch nicht von 7 an rückwärts. Fast ein Viertel der Kinder konnte die Anzahl von 12 abgebildeten Schildkröten nicht durch Zählen ermitteln und 17 % aller Kinder konnten Zahlen nicht von anderen Zeichen unterscheiden. Daneben gab es Kinder, die formale Additions- und Subtraktionsaufgaben im Zwanzigerraum und darüber lösen, Zahlen im Hunderterraum vergleichen und problemlos Relationen zwischen Zahlen herstellen konnten. Insgesamt ist davon auszugehen, dass ca. 20 % der Kinder am Ende des Kindergartens über einen beachtlichen Teil der mathematischen Kompetenzen verfügen, die sie gemäss Lehr- und Bildungsplänen im Verlaufe des ersten Schuljahres zu erwerben haben. Hier sind sowohl die pädagogischen Fachkräfte der Grundschule als auch die mathematischen Schulbuchautoren gefordert, ein Lernangebot zur Verfügung zu stellen, wovon alle Kinder profitieren können – eine Herausforderung für alle.

9.4 Kritische Reflexion der Untersuchung

In der vorliegenden Studie konnten Daten analysiert werden, die in zwei Ländern erhoben worden sind. Die beiden Stichproben in der Schweiz und in Deutschland sind von der Grösse her vergleichbar. Allerdings sind sie für das jeweilige Land nicht repräsentativ. In der Schweiz wurden nur Klassen aus deutschsprachigen Regionen einbezogen. Obwohl Klassen aus verschiedenen Kantonen (St. Gallen, Zürich, Zug, Luzern, Bern, Solothurn und Freiburg) teilgenommen haben, gab es auch Kantone, die in der Stichprobe nicht vertreten waren. In Deutschland waren nur Klassen aus dem Raum Kiel und Vechta in der Untersuchung vertreten.

Aus diesem Grund kann die unterschiedliche mathematische Leistungsentwicklung in den beiden länderspezifischen Stichproben nicht als repräsentativ für das jeweilige Land angesehen werden. Die Ergebnisse geben zwar Hinweise auf mögliche Differenzen, aber es braucht weitere Forschungsarbeiten dazu. Zudem konnten sich in beiden Ländern die pädagogischen Fachkräfte freiwillig zur Teilnahme an der Studie melden. Es handelt sich also um eine sogenannte Selbstselektionsstichprobe (Döring & Bortz, 2016). Entsprechend muss angenommen werden, dass die teilnehmenden pädagogischen Fachkräfte ein besonderes Interesse an mathematischer Bildung im Kindergarten und evtl. auch eine grössere Motivation bezüglich mathematischer Förderung aufweisen, als dies bei Fachkräften der Fall ist, die nicht an der Studie teilgenommen haben. Da aber alle pädagogischen Fachkräfte freiwillig teilnahmen, kann von vergleichbaren motivationalen Aspekten ausgegangen werden.

Weiter wurde in der Studie ersichtlich, dass mit den einbezogenen Variablen zur Erklärung der Unterschiede ohne Einbezug des Vorwissens 34 % und mit Einbezug des Vorwissens 69 % bzw. 71 % der Varianz der Unterschiede in den mathematischen Kompetenzen erklärt werden konnte. Es stellt sich nun die Frage, ob unter Einbezug anderer Variablen eine höhere Erklärungskraft hätte erreicht werden können. Beispielsweise wurden der sozioökonomische Status oder das Unterstützungsverhalten der Eltern nicht erfasst – beides Faktoren, die einen Einfluss auf die mathematischen Kompetenzen der Kinder haben könnten (z. B. Anders et al., 2012; Lehrl et al., 2014; Schuchardt et al., 2014), (vgl. Abschnitt 5.5.1 und 5.5.4).

Eine weitere Limitation der Studie bezieht sich auf das Studiendesign, das keine Gruppe von Kindergärten vorsah, die keine Spiele zur mathematischen Förderung erhielten. Die Ergebnisse zur Leistungsentwicklung der Gesamtstichprobe haben belegt, dass die Kinder zwischen dem zweiten und dritten Testzeitpunkt einen signifikant grösseren mathematischen Kompetenzzuwachs erzielten als zwischen dem ersten und zweiten. In dieser Zeit wurden die Spiele zur mathematischen Förderung eingesetzt. Die Spiele könnten deshalb eine mögliche Erklärung für die unterschiedliche Leistungsentwicklung zwischen den Testzeitpunkten sein. Allerdings gibt es auch andere mögliche Gründe, die zur Erklärung beigezogen werden können (vgl. Abschnitt 9.3.1). Um zu überprüfen, ob die Spiele tatsächlich einen Einfluss auf die mathematische Leistungsentwicklung hatten, wäre ein Vergleich mit einer Kontrollgruppe ohne Spielmaterial notwendig. Die WILMA-Gesamtstudie fokussierte aber andere Fragestellungen (vgl. http://p3.snf.ch/project-156680), sodass im Studiendesign keine Gruppe von Fachkräften vorgesehen war, die keine mathematischen Spiele erhalten hätte.

9.5 Fazit zur Gesamtuntersuchung

Das Ziel dieser Studie bestand einerseits in der Eignungsüberprüfung eines bestehenden mathematischen Testinstrumentes für längsschnittliche Messungen im Kindergarten. Andererseits sollte die mathematische Leistungsentwicklung von Kindergartenkindern über drei Messzeitpunkte analysiert und bedeutsame Einflussfaktoren aufgezeigt werden.

Die Ergebnisse der Überprüfung des adaptierten Tests TEDI-MATH unterstreichen die Wichtigkeit der Auswahl eines adäquaten Testinstruments. Es muss hohe testtheoretische Anforderungen erfüllen. Zum einen müssen genügend Items im niedrigen wie im hohen Leistungsbereich vorhanden sein, um der schnellen Leistungsentwicklung und der grossen Heterogenität von Kindern dieser Altersstufe gerecht zu werden. Nur so kann ein Boden- bzw. Deckeneffekt zu allen drei Testzeitpunkten vermieden werden. Zum anderen sollte der Test neben einer guten Reliabilität auch die Forderung nach Item- und Personenhomogenität erfüllen, um Messfehler möglichst zu minimieren. Entsprechend sollten die Items über gute Fit-Werte verfügen und das Kriterium der Messinvarianz erfüllen, sowohl bezüglich der Testzeitpunkte als auch in Bezug auf Subgruppen. In dieser Untersuchung mussten ca. 40 % der Items wegen schlechtem Item-Fit bzw. Messvarianz aus den Analysen ausgeschlossen werden. Die verbliebenen 46 bzw. 54 Items zeigten sich Rasch-homogen und wiesen zu allen drei Testzeitpunkten eine gute Reliabilität auf. Somit konnten valide Aussagen zur mathematischen Leistungsentwicklung der Kinder gemacht werden. Insgesamt weisen diese Ergebnisse aber darauf hin, dass nicht per se von Itemhomogenität und Messinvarianz ausgegangen werden kann und eine Überprüfung der Items mit dem Rasch-Modell für valide längsschnittliche Analysen bedeutsam ist.

In Bezug auf Unterschiede in den mathematischen Kompetenzen zum ersten Testzeitpunkt konnten die erfassten individuellen Faktoren wie Alter, Geschlecht, Erstsprache und kognitive Fähigkeiten 34 % der Varianz erklären. Unter Einbezug des mathematischen Vorwissens konnte eine Varianzaufklärung von 69 % beim zweiten bzw. 71 % beim dritten Testzeitpunkt erreicht werden, wobei jeweils das Vorwissen die grösste Erklärungskraft besass, gefolgt von den kognitiven Fähigkeiten. Trotz dieser insgesamt hohen Varianzaufklärung bleibt ein gewisser Spielraum für andere mögliche Einflussfaktoren. Aus der Forschung ist bekannt, dass das in dieser Untersuchung nicht erhobene elterliche Unterstützungsverhalten (vgl. Abschnitt 5.5.4) oder der sozioökonomische Status der Familie (vgl. Abschnitt 5.5.1) einen relevanten Einfluss hat. Andere Faktoren wie beispielsweise Aspekte des Unterrichts, der Einfluss von Geschwistern oder Peers sind noch weniger gut erforscht.

Im Gegensatz zu Forschungsarbeiten aus der Grundschule hat in dieser Untersuchung die Leistungsstreuung vom ersten bis zum dritten Testzeitpunkt zu- und nicht abgenommen. Diese Ergebnisse geben einen Hinweis darauf, dass sich im Kontext Kindergarten im Gegensatz zur Grundschule in Bezug auf die mathematische Kompetenzstreuung keine egalisierende Tendenz zeigt. Um diese Tendenz zu bestätigen, braucht es weitere längsschnittliche Forschungsarbeiten im Kindergarten.

Weiter konnte in vorliegender Forschungsarbeit belegt werden, dass das zum ersten Testzeitpunkt erhobene mathematische Vorwissen für die weitere Leistungsentwicklung während der Kindergartenzeit von zentraler Bedeutung ist. Es scheint also auch entscheidend zu sein, welche mathematischen Kompetenzen Kinder bereits vor dem vierten Altersjahr erwerben konnten. Die Ergebnisse des Ländervergleichs geben diesbezüglich einen Hinweis, dass sich möglicherweise die Dauer des Kindergartenaufenthaltes positiv auf die mathematischen Kompetenzen zum ersten Testzeitpunkt ausgewirkt haben könnten (vgl. Abschnitt 9.2.7). Die Kinder in Deutschland hatten zu diesem Zeitpunkt nämlich höhere Kompetenzen als die Kinder in der Schweiz. Während die Kinder in der Schweiz erst ein halbes Jahr im Kindergarten waren, konnten die Kinder in Deutschland die Kindertagesstätte bereits ab dem ersten Lebensjahr besuchen. Von Interesse wäre deshalb eine Untersuchung, ob sich in der Schweiz der Besuch von Einrichtungen im Zusammenhang mit früher Bildung wie Familienzentren, Spielgruppen oder Kindertagesstätten vor dem Kindergarteneintritt positiv auf die mathematischen Kompetenzen auswirkt.

Betrachtet man die mathematische Kompetenzentwicklung über die drei Testzeitpunkte hinweg, wird deutlich, dass die Kinder in der Schweiz einen signifikant höheren Leistungszuwachs und kurz vor dem Schuleintritt auch signifikant höhere mathematische Kompetenzen aufwiesen. Dies wurde so nicht erwartet. Die Frage nach der Begründung dieser Unterschiede kann nicht abschliessend beantwortet werden. Mit Blick auf die Unterschiede der frühen Bildung in den beiden Ländern (vgl. Kapitel 2) handelt es sich hierbei aber nicht primär um Länderunterschiede, sondern um unterschiedliche Kontexte, in denen frühe Bildung stattfindet. Traditionsgemäss gehört der Kindergarten in Deutschland zum Sozialbereich und hat primär einen Betreuungs- und Erziehungsauftrag, während der Kindergarten in der Schweiz schon immer dem Schulwesen nahestand und heute mittlerweile in fast allen Kantonen zur obligatorischen Schulbildung gehört. Entsprechend unterscheiden sich die beiden Länder auch bezüglich der Ausbildung ihrer pädagogischen Fachkräfte. In Deutschland arbeiten mehrheitlich staatlich anerkannte Erzieherinnen und Erzieher in den Kindertagesstätten, die in einer Fachschule für Sozialpädagogik ausgebildet wurden. In der Schweiz erfolgt die Ausbildung zur Kindergärtnerin/zum Kindergärtner auf Tertiärstufe gemeinsam mit den Lehrpersonen der Primarschule an einer pädagogischen Hochschule. Entsprechend ist die Ausbildung in Deutschland in erster Linie sozial-pädagogisch ausgerichtet, während in der Schweiz der Bildungsaspekt stärker im Vordergrund steht. Werden die Unterschiede in der mathematischen Leistungsentwicklung vor dem Hintergrund dieser unterschiedlichen Kontexte betrachtet, scheint sich eine schulnähere, stärker bildungsorientierte Ausrichtung des Kindergartens zumindest im Hinblick auf mathematische Kompetenzen positiv auszuwirken. Es sind allerdings weitere Längsschnittstudien in verschiedenen Kindergarteninstitutionen notwendig, um den Einfluss unterschiedlicher Ausrichtung von Vorschulinstitutionen auf den mathematischen Kompetenzzuwachs der Kinder zu untersuchen.

Literaturverzeichnis

Abt Gürber, N. (2011). *Bereichsspezifische Selbstkonzepte bei Kindern in der Schuleingangsstufe. Zusammenhänge mit Leistungen und Wohlbefinden in der Schule* (Dissertation). Abgerufen von https://doi.org/10.18747/phsg-coll3/id/12

Anders, Y. (2012). *Modelle professioneller Kompetenzen für frühpädagogische Fachkräfte. Aktueller Stand und ihr Bezug zur Professionalisierung. Expertise zum Gutachten «Professionalisierung in der Frühpädagogik»*. München: Knoblingdesign GmbH.

Anders, Y., & Rossbach, H.-G. (2015). Preschool Teachers' Sensitivity to Mathematics in Children's Play: The Influence of Math-Related School Experiences, Emotional Attitudes, and Pedagogical Beliefs. *Journal of Research in Childhood Education, 29*(3), 305–322. https://doi.org/10.1080/02568543.2015.1040564

Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*(2), 231–244. https://doi.org/10.1016/j.ecresq.2011.08.003

Antell, S. E., & Keating, D. P. (1983). Perception of Numerical Invariance in Neonates. *Child Development, 54*(3), 695. https://doi.org/10.2307/1130057

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology, 96*(4), 699–713. https://doi.org/10.1037/0022-0663.96.4.699

Babad, E. Y. (2010). *The social psychology of the classroom*. London: Routledge.

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*(4), 469–520.

Becker, B., & Reimer, D. (Hrsg.). (2010). *Vom Kindergarten bis zur Hochschule: die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, R. (2010). *Soziale Ungleichheit von Bildungschancen und Chancengerechtigkeit – eine Reanalyse mit bildungspolitischen Implikationen. In Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit*. (4. Aufl.). Wiesbaden: Verlag für Sozialwissenschaften.

Benz, C. (2012). Attitudes of Kindergarten Educators about Math. *Journal für Mathematik-Didaktik, 33*(2), 203–232.

Benz, Christiane, Peter-Koop, A., & Grüßing, M. (2015). *Frühe mathematische Bildung: Mathematiklernen der Drei- bis Achtjährigen*. Berlin, Heidelberg: Springer Spektrum.

Berger, A., Tzur, G., & Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences, 103*(33), 12649–12653. https://doi.org/10.1073/pnas.0605350103

Berger, M. (2016). *Geschichte des Kindergartens: von den ersten vorschulischen Einrichtungen des 18. Jahrhunderts bis zur Kindertagesstätte im 21. Jahrhundert*. Frankfurt am Main: Brandes & Apsel.

Blömeke, S., Kaiser, G., Döhrmann, M., & Lehmann, R. (2010). Mathematisches und mathematikdidaktische Wissen angehender Sekundarstufen-I-Lehrkräften im internationalen Vergleich. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. (S. 197–238). Münster: Waxmann.

Bolker, B. (2018). Package 'lme4'. Abgerufen von https://github.com/lme4/lme4/ http://lme4.r-forge.r-project.org/

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: fundamental measurement in the human sciences* (3. Aufl.). New York, London: Routledge, Taylor and Francis Group.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology, 114*(3), 375–388. https://doi.org/10.1016/j.jecp.2012.09.015

Boone, W., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. New York, NY: Springer Berlin Heidelberg.

Bortz, J., & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler: extras online* (7. Aufl.). Berlin Heidelberg: Springer.

Bos, W. (Hrsg.). (2008). *TIMSS 2007: mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183–198). Göttingen: Schwartz.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 37*(1), 62–83. https://doi.org/10.1111/j.2044-8317.1984.tb00789.x

Browne, M. W., & Cudeck, R. (1989). Single Sample Cross-Validation Indices for Covariance Structures. *Multivariate Behavioral Research, 24*(4), 445–455. https://doi.org/10.1207/s15327906mbr2404_4

Bruner, J. S. (1971). Über kognitive Entwicklung. In J. S. Bruner, R. R. Olver & P. M. Greenfield (Hrsg.), *Studien zur kognitiven Entwicklung. Eine kooperative Untersuchung am «Center for Cognitive Studies» der Harvard-Universität* (S. 21–55). Stuttgart: Klett.

Buddenberg, H., Engels, A., Feltrup, C., Hedtfeld, R., Lierse, H., & Müller, A. (2017). *Mathematik, Kerncurriculum für die Grundschule Schuljahrgänge 1-4*. Hannover: unidruck.

Bullheller, S., & Häcker, H. (2002). *Colored Progressive Matrices (CPM)*. Frankfurt am Main: Swets Test Services.

Bullock, M., & Ziegler, A. (1997). *Entwicklung der Intelligenz und des Denkens: Ergebnisse aus dem SCHOLASTIK-Projekt. Entwicklung im Grundschulalter*. Weinheim: Belz.

Bundesrat. (2006). *Volksabstimmung vom 21. Mai 2006. Erläuterungen des Bundesrates. Neuordnung der Verfassungsbestimmungen zur Bildung*. Bern: Bundeskanzlei.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218–261. https://doi.org/10.1037/a0014412

Ceulemans, A., Baten, E., Loeys, T., Hoppenbrouwers, K., Titeca, D., Rousseau, S., & Desoete, A. (2017). The Relative Importance of Parental Numerical Opportunities, Prerequisite Knowledge and Parent Involvement as Predictors for Early Math Achievement in Young Children. *Interdisciplinary Education and Psychology, 1*(1). https://doi.org/10.31532/InterdiscipEducPsychol.1.1.006

Clarke, B., Clarke, D., Grüßing, M., & Peter-Koop, A. (2013). Mathematische Kompetenzen von Vorschulkindern: Ergebnisse eines Ländervergleichs zwischen Australien und Deutschland. *Journal für Mathematik-Didaktik, 29*, 259–286. https://doi.org/10.1007/bf03339064

Clements, D. H. (1999). Subitizing: What is it? Why teach it? *Teaching Children Mathematics, 5*(7), 400–405.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

D-EDK. (2014). Deutschschweizer Erziehungsdirektoren-Konferenz: Lehrplan 21. Abgerufen von https://www.lehrplan.ch/

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*(1), 83–120.

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study: Mathematical learning disabilities in kindergarten. *British Journal of Educational Psychology, 82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

DESTATIS. (2017). Destatis: Statistisches Bundesamt. Abgerufen von https://www.destatis.de/DE/Startseite.html

Deutscher Bildungsrat. (1970). *Empfehlungen der Bildungskommission. Strukturplan für das Bildungswesen*. Stuttgart: Ernst Klett Verlag.

Ditton, H., & Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *Journal für Bildungsforschung Online, 1*(1), 33–61.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin, Heidelberg: Springer.

Dornheim, D. (2008). *Prädiktion von Rechenleistung und Rechenschwäche: der Beitrag von Zahlen-Vorwissen und allgemein-kognitiven Fähigkeiten*. Berlin: Logos-Verl.

Dorsch, F., Wirtz, M. A., & Strohmer, J. (Hrsg.). (2017). *Dorsch - Lexikon der Psychologie* (18. Aufl.). Bern: Hogrefe.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Dunekacke, S., Jenßen, L., & Blömeke, S. (2015). Effects of Mathematics Content Knowledge on Pre-school Teachers' Performance: a Video-Based Assessment of Perception and Planning Abilities in Informal Learning Situations. *International Journal of Science and Mathematics Education, 13*(2), 267–286. https://doi.org/10.1007/s10763-014-9596-z

EDK. (2003). Schweizerische Konferenz der kantonalen Erziehungsdirektoren: Aktionsplan „Pisa 2000"- Folgemassnahmen. Beschluss Plenarversammlung 12. Juni 2003. Abgerufen von http://www.edk.ch/dyn/25602.php

EDK. (2007). Schweizerische Konferenz der kantonalen Erziehungsdirektoren: Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule (HarmoS-Konkordat). Abgerufen von http://www.edk.ch/dyn/11659.php

EDK. (2017). Schweizerische Konferenz der kantonalen Erziehungsdirektoren: Studentafeln der Volksschule. Abgerufen von https://edudoc.ch/collection/Stundentafel?ln=de

Ehm, J.-H., Duzy, D., & Hasselhorn, M. (2011). Das akademische Selbstkonzept bei Schulanfängern: Spielen Geschlecht und Migrationshintergrund eine Rolle?. *Frühe Bildung*, 37–45. https://doi.org/10.1026/2191-9186/a000008

Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden: mit Online-Materialien* (5. Aufl.). Weinheim Basel: Beltz.

Ennemoser, M., Sinner, D., & Krajewski, K. (2015). Kurz- und langfristige Effekte einer entwicklungsorientierten Mathematikförderung bei Erstklässlern mit drohender Rechenschwäche. *Lernen und Lernstörungen, 4*(1), 43–59. https://doi.org/10.1024/2235-0977/a000091

Erning, G., Neumann, K., & Reyer, J. (1987). *Geschichte des Kindergartens. (Bd. 2: Institutionelle Aspekte systematische Perspektiven, Entwicklungsverläufe)*. Freiburg: Lambertus.

Esser, G., & Wyschkon, A. (2016). *BUEVA-III: Basisdiagnostik Umschriebener Entwicklungsstörungen im Vorschulalter – Version III*. Göttingen: Hogrefe.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' Discrimination of Number vs. Continuous Extent. *Cognitive Psychology, 44*(1), 33–66. https://doi.org/10.1006/cogp.2001.0760

Fischer, U., Roesch, S., & Moeller, K. (2017). Diagnostik und Förderung bei Rechenschwäche: Messen wir, was wir fördern wollen? *Lernen und Lernstörungen, 6*(1), 25–38. https://doi.org/10.1024/2235-0977/a000160

Florida Department of Education. (2004). *Assessment and accountability briefing book*. Tallahasse: Author.

Frey, A., Gehrlein, B., & Wosnitza, M. (2006). *Friedrich Fröbel und seine Pädagogik* (2. Aufl). Landau: Verl. Empirische Pädagogik.

Fritz, A., & Ricken, G. (2008). *Rechenschwäche*. München: Reinhardt.

Fritz, A., Ehlert, A., & Leutner, D. (2018). Arithmetische Konzepte aus kognitiv-entwicklungspsychologischer Sicht. *Journal für Mathematik-Didaktik, 39*(1), 7–41. https://doi.org/10.1007/s13138-018-0131-6

Fthenakis, W. E. (2009a). Bildung neu definieren und hohe Bildungsqualität von Anfang an sichern. Ein Plädoyer für die Stärkung von prozessualer Qualität., *Betrifft Kinder* (3), 6–10. Weimar: Verl. Das Netz.

Fthenakis, W. E. (2009b). *Frühe mathematische Bildung: Kinder unter 6 Jahren*. Troisdorf: Bildunsgverl. EINS.

Fuson, K. C. (1988). *Children's Counting and Concept of Number*. New York: Springer.

Gaidoschik, M. (2010). *Wie Kinder rechnen lernen--oder auch nicht: eine empirische Studie zur Entwicklung von Rechenstrategien im ersten Schuljahr*. Frankfurt am Main: Peter Lang.

Gallit, F., Wyschkon, A., Poltz, N., Moraske, S., Kucian, K., von Aster, M., & Esser, G. (2018). Henne oder Ei: Reziprozität mathematischer Vorläufer und Vorhersage des Rechnens. *Lernen und Lernstörungen, 7*(2), 81–92. https://doi.org/10.1024/2235-0977/a000205

Gasteiger, H. (2010). *Elementare mathematische Bildung im Alltag der Kindertagesstätte: Grundlegung und Evaluation eines kompetenzorientierten Förderansatzes*. Münster, New York, München, Berlin: Waxmann.

Gasteiger, H., Brunner, E., & Chen, C.-S. (2018). Frühe mathematische Bildung in Deutschland, Taiwan und der Schweiz - ein Vergleich der Ausgangslagen. In *Beiträge zum Mathematikunterricht 2018* (S. 587–590). Münster: WTM-Verlag.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLoS ONE, 8*(1), e54651. https://doi.org/10.1371/journal.pone.0054651

Geiser, C. (2011). *Datenanalyse mit Mplus: eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag.

Gelman, R. (1997). Constructing and using conceptual competence. *Cognitive Development, 12*(3), 305–313. https://doi.org/10.1016/S0885-2014(97)90002-2

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number*. Cambridge, Mass: Harvard University Press.

Gerlach, M. (2007). *Entwicklungsaspekte des Rechnenlernens. Fördermöglichkeiten bei beeinträchtigtem Erwerb mathematischer Kompetenzen im Grundschulalter*. Duisburg-Essen.

Gerlach, M., Fritz, A., Ricken, G., & Schmidt, S. (2007). *Baustein 3: Nicht-zählende Rechenstrategien*. Berlin: Cornelsen.

Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the «number sense»: The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*(5), 1457–1465. https://doi.org/10.1037/a0012682

Hartig, J., & Frey, A. (2013). Sind Modelle der Item-Response-Theorie (IRT) das „Mittel der Wahl" für die Modellierung von Kompetenzen? *Zeitschrift für Erziehungswissenschaft, 16*(S1), 47–51. https://doi.org/10.1007/s11618-013-0386-0

Hartig, J., & Kühnbach, O. (2006). Schätzung von Veränderung mit „plausible values" in mehrdimensionalen Rasch-Modellen. In A. Ittel & H. Merkens (Hrsg.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (S. 27–44). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90502-0_3

Hasemann, K., & Gasteiger, H. (2014). *Anfangsunterricht Mathematik* (3. Aufl.). Berlin, Heidelberg: Springer Spektrum.

Hattie, J., Beywl, W., & Zierer, K. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider-Verl. Hohengehren.

Hauser, B., Rathgeb-Schnierer, E., Stebler, R., & Vogt, F. (Hrsg.). (2015). *Mehr ist mehr: mathematische Frühförderung mit Regelspielen*. Seelze: Klett/Kallmeyer.

Hauser, B., Vogt, F., Stebler, R., & Rechsteiner, K. (2014). Förderung früher mathematischer Kompetenzen: Spielintegriert oder trainingsbasiert. *Frühe Bildung, 3*(3), 139–145. https://doi.org/10.1026/2191-9186/a000144

Heinze, A., Herwartz-Emden, L., & Reiss, K. (2007). Mathematikkenntnisse und sprachliche Kompetenz bei Kindern mit Migrationshintergrund zu Beginn der Grundschulzeit. *Zeitschrift für Pädagogik, 53*(4), 562–581.

Hemmerling, A. (2007). *Der Kindergarten als Bildungsinstitution. Hintergründe und Perspektiven*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hess, K. (2012). *Kinder brauchen Strategien: eine frühe Sicht auf mathematisches Verstehen*. Seelze: Kallmeyer.

Hildenbrand, C. (2016). *Förderung früher mathematischer Kompetenzen: eine Interventionsstudie zu den Effekten unterschiedlicher Förderkonzepte*. Münster, New York: Waxmann.

Holder, K., & Kessels, U. (2017). Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective. *Social Psychology of Education, 20*(3), 471–490. https://doi.org/10.1007/s11218-017-9384-z

Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014). Predicting first-grade mathematics achievement: the contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in Psychology, 5*, 272. https://doi.org/10.3389/fpsyg.2014.00272

Hox, J. J., Moerbeek, M., & Schoot, R. van de. (2017). *Multilevel analysis: techniques and applications* (3. Aufl.). New York, NY: Routledge.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Hrsg.), *Structural equation modeling: Concepts, issues, and applications* (S. 76–99). London: Sage.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huntsinger, C. S., Jose, P. E., Liaw, F.-R., & Ching, W.-D. (1997). Cultural Differences in Early Mathematics Learning: A Comparison of Euro-American, Chinese-American, and Taiwan-Chinese Families. *International Journal of Behavioral Development, 21*(2), 371–388. https://doi.org/10.1080/016502597384929

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender Similarities *Characterize Math Performance*. *Science, 321*(5888), 494–495. https://doi.org/10.1126/science.1160364

Ikaheimo, H. (1996). *Diagnostic of the central mathematical concepts*. Helsinki: Oy Opperi Ab.

Ittel, A., & Merkens, H. (Hrsg.). (2006). *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, 106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in Children's Self-Competence and Values: Gender and Domain Differences across Grades One through Twelve. *Child Development, 73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jansen, R. (2010). *Die Ausbildung frühpädagogischer Fachkräfte an Berufsfachschulen und Fachschulen: eine Analyse im Ländervergleich; Expertise für das Projekt Weiterbildungsinitiative Frühpädagogischer Fachkräfte* (WiFF) (2. Aufl.). München: DJI.

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research & Practice, 22*(1), 36–46.

Jörns, C., Schuchardt, K., Grube, D., & Mähler, C. (2014). Spielorientierte Förderung numerischer Kompetenzen im Vorschulalter und deren Eignung zur Prävention von Rechenschwierigkeiten. *Empirische Sonderpädagogik, 3*, 243–259.

Jörns, C., Schuchardt, K., Mähler, C., & Grube, D. (2013). Alltagsintegrierte Förderung numerischer Kompetenzen im Kindergarten. *Frühe Bildung, 2*(2), 84–91. https://doi.org/10.1026/2191-9186/a000088

Jugendministerkonferenz. (2004). Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen. Beschluss vom 13./14.05.2004. Abgerufen von https://www.kmk.org/.../2004/2004_06_03-Fruehe-Bildung-Kindertageseinrichtungen

Karst, K., & Lipowsky, F. (2013). Leistungsentwicklung im Fach Mathematik und deren Determinanten. In F. Lipkowsky, G. Faust & C. Kastens (Hrsg.), *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren* (S. 51–76). Münster, New York, München, Berlin: Waxmann.

Kaufmann, L., Nürk, H. C., Graf, M., Krinzinger, H., Delazer, M., & Hinckeldey, K. (2009). *TEDI-MATH: Test zur Erfassung numerisch-rechnerischer Fertigkeiten vom Kindergarten bis zur 3. Klasse. Deutschsprachige Adaptation des Test diagnostique des compétences de base en mathématiques (TEDI-MATH) von M.P. Noël, J. Grégoire und C. van Nieuwenhoven*. München: Huber.

Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik, 52*(6), 876–903.

Knauer, R., & Hansen, R. (2012). Erfolgreich starten. Leitlinien zum Bildungsauftrag in Kindertagesstätten. Abgerufen von https://www.schleswig-holstein.de/DE/Fachinhalte/K/.../leitlinien_bildungsauftrag.html

Knievel, I., Lindmeier, A. M., & Heinze, A. (2015). Beyond Knowledge: Measuring Primary Teachers' Subject-Specific Competences in and for Teaching Mathematics with Items Based on Video Vignettes. *International Journal of Science and Mathematics Education, 13*(2), 309–329. https://doi.org/10.1007/s10763-014-9608-z

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical Transformations in Five-month-old Human Infants. *Mathematical Cognition, 3*(2), 89–104.

Koglin, U., Janke, N., & Petermann, F. (2009). Werden IQ-Veränderungen vom Kindergarten- zum Schulalter durch psychosoziale Risikofaktoren beeinflusst? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41*(3), 132–141. https://doi.org/10.1026/0049-8637.41.3.132

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95–103. https://doi.org/10.1016/j.learninstruc.2012.12.001

Koller, I., Alexandrowicz, R., & Hatzinger, R. (2012). *Das Rasch-Modell in der Praxis: eine Einführung mit eRm*. Wien: facultas.wuv.

Konrad, F.-M. (2004). *Der Kindergarten: seine Geschichte von den Anfängen bis in die Gegenwart*. Freiburg im Breisgau: Lambertus.

Krajewski, K. & Ennemoser, M. (2013). Entwicklung und Diagnostik der Zahl-Größen-Verknüpfung zwischen 3 und 8 Jahren. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Hrsg.), *Diagnostik mathematischer Kompetenzen. Tests & Trends N.F. 11* (S. 41–65). Göttingen: Hogrefe.

Krajewski, K. (2008). *Vorhersage von Rechenschwäche in der Grundschule* (2. Aufl.). Hamburg: Verlag Dr. Kovač.

Krajewski, K. (2018). *MBK O. Test Mathematischer Basiskompetenzen im Kindergartenalter*. Göttingen: Hogrefe.

Krajewski, K., & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht, 53*, 246–262.

Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction, 19*(6), 513–526. https://doi.org/10.1016/j.learninstruc.2008.10.002

Krajewski, K., Küspert, P., & Schneider, W. (2002). *Deutscher Mathematiktest für erste Klassen (DEMAT 1 +)*. Göttingen: Hogrefe.

Krajewski, K., Renner, A., Nieding, G., & Schneider, W. (2009). Frühe Förderung von mathematischen Kompetenzen im Vorschulalter. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (S. 91–103). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91452-7_7

Kristen, C., & Granato, N. (2007). Bildungsinvestitionen in Migrantenfamilien. In Bundesministerium für Bildung und Forschung (Hrsg.), *Migrationshintergrund von Kindern und Jugendlichen: Wege zur Weiterentwicklung der amtlichen Statistik* (S. 25–42). Bonn, Berlin: BMBF.

Kucharz, D. (Hrsg.). (2014). *Professionelles Handeln im Elementarbereich (PRIMEL): eine deutsch-schweizerische Videostudie*. Münster, New York: Waxmann.

Kucian, K., von Aster, M., Loenneker, T., Dietrich, T., & Martin, E. (2008). Development of Neural Networks for Exact and Approximate Calculation: A fMRI Study. *Developmental Neuropsychology, 33*(4), 447–473. https://doi.org/10.1080/87565640802101474

Langhorst, P., Hildenbrand, C., Ehlert, A., & Fritz, A. (2013). Mathematische Bildung im Kindergarten – Evaluation des Förderprogramms «Mina und der Maulwurf» und Betrachtung von Fortbildungsvarianten. In M. Hasselhorn, A. Heinze, W. Schneider, & U. Trautwein (Hrsg.), *Tests und Trends. Jahrbuch der pädagogisch-psychologischen Diagnostik. (Bd. 11)*. Göttingen: Hogrefe.

Lehnemann, W. (1995). Öffentliche Kleinkindererziehung im 19. Jahrhundert: Der erste Kindergarten in Westfalen. Abgerufen von froebel- museum.de/media/.books/pdf_95. pdf

Lehrl, S., Kuger, S., & Anders, Y. (2014). Soziale Disparitäten beim Zugang zu Kindergartenqualität und differenzielle Konsequenzen für die vorschulische mathematische Entwicklung. *Unterrichtswissenschaft, 42*(2), 132–151.

Levin, A., Wittmann, G., & Bönig, D. (2016). *Mathematikbezogene Überzeugungen. In AnschlussM. Anschlussfähigkeit mathematikdidaktischer Überzeugungen und Praktiken von ErzieherInnen und GrundschullehrerInnen*. Münster: Waxmann.

Li, X., Liu, S., DeBey, M., McFadden, K., & Pan, Y.-J. (2018). Investigating Chinese preschool teachers' beliefs in mathematics teaching from a cross-cultural perspective. *Early Years, 38*(1), 86–101. https://doi.org/10.1080/09575146.2016.1228615

Link, M., Vogt, F., & Hauser, B. (2017). Überzeugungen von Kindergartenlehrpersonen zur mathematischen Förderung im Kindergarten: Die Schweiz, Deutschland und Österreich im Vergleich. *Beiträge zur Lehrerinnen- und Lehrerbildung, 35*(3), 440–458.

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science, 20*(4), e12418. https://doi.org/10.1111/desc.12418

Lyons, I. M., Bugden, S., Zheng, S., De Jesus, S., & Ansari, D. (2018). Symbolic number skills predict growth in nonsymbolic number skills in kindergarteners. *Developmental Psychology, 54*(3), 440–457. https://doi.org/10.1037/dev0000445

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science, 17*(5), 714–726. https://doi.org/10.1111/desc.12152

Manfra, L., Squires, C., Dinehart, L. H. B., Bleiker, C., Hartman, S. C., & Winsler, A. (2017). Preschool writing and premathematics predict Grade 3 achievement for low-income, ethnically diverse children. *The Journal of Educational Research, 110*(5), 528–537. https://doi.org/10.1080/00220671.2016.1145095

Mantzicopoulos, P. (2006). Younger Children's Changing Self-Concepts: Boys and Girls From Preschool Through Second Grade. *The Journal of Genetic Psychology, 167*(3), 289–308. https://doi.org/10.3200/GNTP.167.3.289-308

Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology, 38*(3), 376–393.

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School. *Journal of Social Issues, 64*(1), 95–114. https://doi.org/10.1111/j.1540-4560.2008.00550.x

Metzinger, A. (1993). *Zur Geschichte der Erzieherausbildung: Quellen, Konzeptionen, Impulse, Innovationen*. Frankfurt am Main, New York: P. Lang.

Ministerium für Bildung und Wissenschaft des Landes Schleswig-Holstein. (2013). Lehrplan für die Fachschule. Fachrichtung Sozialpädagogik. Ausbildungsgang Erzieherin/Erzieher. Abgerufen von https://lehrplan.lernnetz.de/index.php?wahl=156

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: another look. *Developmental Psychology, 33*(3), 423–428.

Moore, D., Benenson, J., Reznick, J. S., Peterson, M., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology, 23*(5), 665–670. https://doi.org/10.1037/0012-1649.23.5.665

Moosbrugger, H., & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 7–26). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-20072-4_2

Moser Opitz, E. (1999). Mathematischer Erstunterricht im heilpädagogischen Bereich: Anfragen und Überlegungen. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete, 68*(3), 293–307.

Moser Opitz, E. (2008). *Zählen, Zahlbegriff, Rechnen: theoretische Grundlagen und eine empirische Untersuchung zum mathematischen Erstunterricht in Sonderklassen* (3. Aufl.). Bern, Stuttgart, Wien: Haupt Verlag.

Moser Opitz, E. (2012). Erstrechen. In Didaktik des Unterrichts im Förderschwerpunkt Lernen. Ein Handbuch für Studium und Praxis. (2. Aufl.). Stuttgart: Kohlhammer.

Moser Opitz, E. (2013). *Rechenschwäche/Dyskalkulie: theoretische Klärungen und empirische Studien an betroffenen Schülerinnen und Schülern* (2. Aufl.). Bern: Haupt.

Moser Opitz, E., Ruggiero, D., & Wüest, P. (2010). Verbale Zählkompetenzen und Mehrsprachigkeit: Eine Studie mit Kindergartenkindern. *Psychologie in Erziehung und Unterricht, 57*(3), 161–174. https://doi.org/10.2378/peu2010.art12d

Moser, U., Stamm, M., & Hollenweger, J. (2004). *Lernstandserhebung 1. Klassen Kanton Zürich. Schlussbericht.* Zürich: Bildungsdirektion.

Nachtigall, C., & Suhl, U. (2002). *Der Regressionseffekt. Mythos und Wirklichkeit. methevalreport 4(2).* Abgerufen von www.uni-jena.de/svw/metheval/report/

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R 2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Niedersächsisches Kultusministerium. (2011). Orientierungsplan für Bildung und Erziehung im Elementarbereich niedersächsischer Tageseinrichtungen für Kinder. Abgerufen von https://www.mk.niedersachsen.de/download/4491

Niklas, F., & Schneider, W. (2012). Die Anfänge geschlechtsspezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 44*(3), 123–138. https://doi.org/10.1026/0049-8637/a000064

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology, 3*(1), 1–18. https://doi.org/10.1016/0022-2496(66)90002-2

Nufer, H. (2012). Kindergarten. Historisches Lexikon der Schweiz. Abgerufen von http://www.hls-dhs-dss.ch/textes/d/D10401.php

Oberhuemer, P., Schreyer, I., & Neuman, M. J. (2010). *Professionals in early childhood education and care systems: European profiles and perspectives*. Opladen, Farmington Hills, MI: Verlag Barbara Budrich.

Obersteiner, A. (2012). *Mentale Repräsentationen von Zahlen und der Erwerb arithmetischer Fähigkeiten: Konzeptionierung einer Förderung mit psychologisch-didaktischer Grundlegung und Evaluation im ersten Schuljahr*. Münster: Waxmann.

Padberg, F., & Benz, C. (2011). *Didaktik der Arithmetik: für Lehrerausbildung und Lehrerfortbildung* (4. Aufl.). Heidelberg: Spektrum Akademischer Verlag.

Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research, 62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Paek, I., & Wilson, M. (2011). Formulating the Rasch Differential Item Functioning Model Under the Marginal Maximum Likelihood Estimation Context and Its Comparison With Mantel–Haenszel Procedure in Short Test and Small Sample Conditions. *Educational and Psychological Measurement*, *71*(6), 1023–1046. https://doi.org/10.1177/0013164411400734

Park, M. (2000). Linguistic Influence on Numerical Development. *The Mathematics Educator, 10*(1), 19–24.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*(1), 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

PHSG. (o.D.). Pädagogische Hochschule St.Gallen. Studium Kindergarten und Primarschule. Studienorganisation. Modulbeschreibungen HeS18 und FrS19. Abgerufen von http://www.extranet.phsg.ch/extranet/studium-kindergarten-und-primarschule/studienorganisation/modulbeschreibungen.aspx

Post, S., Kastens, C., & Lipowsky, F. (2013). Professionelle Handlungskompetenz von Lehrpersonen. In C. Kastens (Hrsg.), *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beide*, New York, München, Berlin: Waxmann.

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2015). Sprachkompetenz und Mathematikleistung – Empirische Untersuchung sprachlich bedingter Hürden in den Zentralen Prüfungen 10. *Journal für Mathematik-Didaktik, 36*(1), 77–104. https://doi.org/10.1007/s13138-015-0074-0

Prenzel, M., Heidemeier, H., Ramm, G., Hohensee, F., & Ehmke, T. (2004). Soziale Herkunft und mathematische Kompetenz. In PISA-Konsortium Deutschland (Hrsg.*), Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 273–282). Münster: Waxmann.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Radatz, H., & Schipper, W. (1983). *Handbuch für den Mathematikunterricht an Grundschulen*. Hannover: Schrödel.

Radatz, H., Schipper, W., Ebeling, A., & Dröge, R. (Hrsg.). (1996). *Handbuch für den Mathematikunterricht. 1. Schuljahr*. Hannover: Schroedel.

Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*. Kopenhagen: The Danish Institute for Education Research.

RBZ1. (o.D.). Regionales Berufsbildungszentrum. Soziales, Ernährung und Bau der Landeshauptstadt Kiel. Abgerufen von https://www.rbz1.de/erzieher-in.html

Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E., & Köller, O. (Hrsg.). (2016). *PISA 2015: eine Studie zwischen Kontinuität und Innovation*. Münster, New York: Waxmann.

Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist, 44*(2), 162–169. https://doi.org/10.1037/0003-066X.44.2.162

Ricken, G., Fritz, A., & Balzer, L. (2013). *MARKO-D: Mathematik- und Rechenkonzepte im Vorschulalter - Diagnose (Hogrefe Vorschultests)*. Göttingen: Hogrefe.

Rindermann, H., & Baumeister, A. E. E. (2015). Parents' SES vs. parental educational behavior and children's development: A reanalysis of the Hart and Risley study. *Learning and Individual Differences, 37*, 133–138. https://doi.org/10.1016/j.lindif.2014.12.005

Robitzsch, A., Kiefer, T., & Wu, M. (2018). *Package TAM, Test Analysis Modules*. R package version 3.0-21. https://CRAN

Rohe, A. M., & Quaiser-Pohl, C. (2010). Prädiktoren für mathematische Kompetenzen zu Beginn der Grundschule - Gibt es Unterschiede zwischen Mädchen und Jungen? In C. Quaiser-Pohl & M. Endepohls-Ulpe (Hrsg.), *Bildungsprozesse im MINT-Bereich. Interesse, Partizipation und Leistungen von Mädchen und Jungen* (S. 13–27). Münster: Waxmann.

Rösch, H., & Paetsch, J. (2011). Sach- und Textaufgaben im Mathematikunterricht als Herausforderung für mehrsprachige Kinder. In S. Prediger & E. Özdil (Hrsg.), *Mathematiklernen unter Bedingungen der Mehrsprachigkeit* (S. 55–76). Münster: Waxmann.

Rosseel, Y. (2017). Package ‘lavaan’. Abgerufen von https://cran.r- project.org/web/packages /lavaan/lavaan.pdf.

Rost, D. H. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien: eine Einführung* (3. Aufl.). Bad Heilbrunn: Verlag Julius Klinkhardt.

Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2. Aufl.). Bern: Huber.

Sahr, K. (2012). *Geschlechtsunterschiede in mathematischen Fähigkeiten und akademischen Selbstkonzepten bei Kindern im Vor- und frühen Grundschulalter* (Dissertation). Duisburg-Essen: Universität Duisburg-Essen.

Sale, A., Schell, A., Koglin, U., & Hillenbrand, C. (2018). Einflussfaktoren mathematischer Kompetenzen vor Schuleintritt. *Empirische Sonderpädagogik*, *4*, 370–387.

Scherer, P., & Moser Opitz, E. (2010). *Fördern im Mathematikunterricht der Primarstufe*. Heidelberg: Spektrum Akademischer Verlag.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Schilling, S. R., Sparfeldt, J. R., & Rost, D. H. (2006). Facetten schulischen Selbstkonzepts: Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie, 20*(1/2), 9–18. https://doi.org/10.1024/1010-0652.20.12.9

Schlegel-Ganz, N. (2002). *Die Entwicklung des Kindergartens im Kanton St.Gallen. In Beiträge zur ostschweizerischen Schulgeschichte*. St.Gallen: Historischer Verein des Kantons St.Gallens.

Schmidt, R. (1982). *Zahlenkenntnisse von Schulanfängern. Ergebnisse einer Untersuchung, Sachunterricht und Mathematik in der Primarschule*. Wiesbaden: Hessisches Institut für Bildungsplanung und Schulentwicklung.

Schmiedek, F., & Wolff, J. K. (2010). Latente Wachstumskurvenmodelle. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 1017–1029). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_38

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350–353.

Schneider, W., Küspert, P., & Krajewski, K. (2016). *Die Entwicklung mathematischer Kompetenzen* (2. Aufl.). Paderborn: Ferdinand Schöningh.

Schnell, R., Hill, P., & Esser, E. (1999). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

Schuchardt, K., Piekny, J., Grube, D., & Mähler, C. (2014). Einfluss kognitiver Merkmale und häuslicher Umgebung auf die Entwicklung numerischer Kompetenzen im Vorschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 46*(1), 24–34. https://doi.org/10.1026/0049-8637/a000099

Schuh-Custer, A. (1969). Die Ausbildung der Kindergärtnerinnen und Hortnerinnen in der deutschsprachigen Schweiz. *Archiv für das schweizerische Unterrichtswesen, 54/55*, 131–137. https://doi.org/10.5169/seals-59456

Schuntermann, M. (2001). International Classification of Functioning, Disability and Health (ICF) by WHO - Short Summary. Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin, 11(6), 229–230. https://doi.org/10.1055/s-2001-19074

Schwab, S., & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik, 7*(3), 175–193.

Seeger, D., Holodynski, M., & Roth, M. (2018). BIKO-Mathekiste: Spielbasierte Förderung für 4- bis 6-jährige Kinder mit einem Entwicklungsrisiko im Bereich numerischer Basiskompetenz. *Frühe Bildung, 7*(3), 135–143. https://doi.org/10.1026/2191-9186/a000380

Selter, C., & Spiegel, H. (2007). *Wie Kinder rechnen*. Leipzig: Klett-Grundschulverl.

Shulman, L. (1986). Those Who Unterstand: Knowledge Growth in Teaching. *Educational Researcher, 15*(2), 4–14.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560. https://doi.org/10.1037/a0014239

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10*(2), 253–269. https://doi.org/10.1016/0885-2014(95)90011-X

Smith, W. (2013). Vorschulische Förderung im Kindergarten. In G. Faust (Hrsg.), *Einschulung: Ergebnisse aus der Studie Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)* (S. 69–82). Münster: Waxmann.

Smith, E. V., & Smith, R. M. (2004). *Introduction to Rasch measurement: Theory, models and applications*. Maple Grove, Minn: JAM Press.

Stamm, M. (2005). *Zwischen Exzellenz und Versagen: Frühleser und Frührechnerinnen werden erwachsen*. Zürich: Rüegger.

Starkey, P, & Cooper, R. (1980). Perception of numbers by human infants. *Science, 210*(4473), 1033–1035. https://doi.org/10.1126/science.7434014

Stern, E. (1998). *Die Entwicklung des mathematischen Verständnisses im Kindesalter*. Lengerich: Pabst.

Stern, E., & Neubauer, A. (2013). *Intelligenz: große Unterschiede und ihre Folgen* (2. Aufl). München: Dt. Verl.-Anst.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development, 52*(4), 1146–1152.

Strobl, C. (2015). *Das Rasch-Modell: eine verständliche Einführung für Studium und Praxis* (3. Aufl.). München: Hampp.

Szkudlarek, E., & Brannon, E. M. (2017). Does the Approximate Number System Serve as a Foundation for Symbolic Mathematics? *Language Learning and Development, 13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Tenorth, H.-E. (2010). *Geschichte der Erziehung: Einführung in die Grundzüge ihrer neuzeitlichen Entwicklung* (5. Aufl.). Weinheim, München: Juventa Verlag.

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics, 41*(2), 191–2017.

Tiedemann, J., & Billmann-Mahecha, E. (2004). Kontextfaktoren der Schulleistung im Grundschulalter: Ergebnisse aus der Hannoverschen Grundschulstudie. *Zeitschrift für Pädagogische Psychologie, 18*(2), 113–124. https://doi.org/10.1024/1010-0652.18.2.113

Toll, S. W. M., Kroesbergen, E. H., & Van Luit, J. E. H. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *British Journal of Educational Psychology, 86*(3), 429–445. https://doi.org/10.1111/bjep.12116

Ufer, S., & Mehringer, V. (2013). Sprachstand, soziale Herkunft und Bilingualität: Effekte auf Facetten mathematischer Kompetenz. In M. Becker-Mrotzk, K. Schramm, E. Thürmann, & H. J. Vollmer (Hrsg.), *Sprache im Fach – Sprachlichkeit und fachliches Lernen* (S. 167–184). Münster: Waxmann.

Valeski, T. N., & Stipek, D. J. (2001). Young Children's Feelings about School. *Child Development, 72*(4), 1198–1213. https://doi.org/10.1111/1467-8624.00342

von Aster, M. G. (2001). *Die Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI)*. Frankfurt am Main: Harcourt.

von Aster, M., Bzufka, M. W., & Horn, R. R. (2009). *Neurologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern – Kindergartenversion – (ZAREKI–K)*. Frankfurt am Main: Pearson.

von Aster, M., Schweiter, M., & Weinhold Zulauf, M. (2007). Rechenstörungen bei Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39*(2), 85–96. https://doi.org/10.1026/0049-8637.39.2.85

von Aster, M. G, & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49*(11), 868–873. https://doi.org/10.1111/j.1469-8749.2007.00868.x

Walter-Laager, C., Joos, Y., & Fröhlich, W. (Hrsg.). (2013). *Kindergarten: Grundlagen aktueller Kindergartendidaktik* (3. Aufl.). Winterthur: ProKiga-Lehrmittelverl.

Wannack, E. (2003). Kindergarten und Schule - Lehrpläne im Vergleich. *Schweizerische Zeitschrift für Bildungswissenschaften, 25*(2), 271–286.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450. https://doi.org/10.1007/BF02294627

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's Past Is Prologue: Relations Between Early Mathematics Knowledge and High School Achievement. *Educational Researcher, 43*(7), 352–360. https://doi.org/10.3102/0013189X14553660

Weinert, F. E. (Hrsg.). (2001). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–31). Weinheim: Beltz.

Weinert, Franz Emanuel, & Helmke, A. (Hrsg.). (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz Psychologie-Verl.-Union.

Weinhold Zulauf, M., Schweiter, M., & von Aster, M. (2003). Das Kindergartenalter: Sensitive Periode für die Entwicklung numerischer Fertigkeiten. *Kindheit und Entwicklung, 12*(4), 222–230. https://doi.org/10.1026//0942-5403.12.4.222

Weise, G. (1975). *Psychologische Leistungstests*. Göttingen: Hogrefe.

Weiss, R. H., & Osterland, J. (2012). *Grundintelligenztest Skala 1 - Revision. CFT 1-R*. Göttingen: Hogrefe.

Weiss, R. H., Cattell, R. B., & Osterland, J. (1997). *CFT 1. Grundintelligenztest Skala 1*. Göttingen: Hogrefe.

Weisshaupt, S., & Peucker, S. (2009). Entwicklung arithmetischen Vorwissens. In A. Fritz, G. Ricken & S. Schmidt (Hrsg.), *Handbuch Rechenschwäche* (S. 52–76). Weinheim: Beltz.

Weisshaupt, S., Peucker, S., & Wirtz, M. A. (2006). Diagnose mathematischen Vorwissens im Vorschulalter und Vorhersage von Rechenleistungen und Rechenschwierigkeiten in der Grundschule. *Psychologie in Erziehung und Unterricht, 53*, 236–245.

Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K., & Kasper, D. (Hrsg.). (2016). *TIMSS 2015: mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster, New York: Waxmann.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology, 44*(2), 588–596. https://doi.org/10.1037/0012-1649.44.2.588

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial Invariance Within Longitudinal Structural Equation Models: Measuring the Same Construct Across Time. Child *Development Perspectives, 4*(1), 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, N.J: Lawrence Erlbaum Associates.

Winkelmann, H., Heuvel-Panhuizen, M., & Robitzsch, A. (2008). Gender differences in the mathematics achievements of German primary school students: results from a German large-scale study. *ZDM, 40*(4), 601–616. https://doi.org/10.1007/s11858-008-0124-x

Wittmann, E. C. (2010). Grundsätzliche Überlegungen zur frühkindlichen Bildung in der Mathematik. In M. Stamm & D. Edelmann (Hrsg.), *Frühkindliche Bildung, Betreuung und Erziehung*. (S. 177–195). Zürich: Rüegger.

Witzig, H. (2013). Geschichte des Kindergartens. In C. Walter & K. Fasseing (Hrsg.), *Grundlagen aktueller Kindergartenpädagogik* (3. Aufl.). Winterthur: ProKiga-Lehrmittelverlag.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock- Johnson tests of achievement*. Itasca: Riverside.

Wullschleger, A. (2017). *Individuell-adaptive Lernunterstützung im Kindergarten: eine Videoanalyse zur spielintegrierten Förderung von Mengen-Zahlen-Kompetenzen*. Münster, New York: Waxmann.

Wynn, K. (1996). Infants' Individuation and Enumeration of Actions. *Psychological Science, 7*(3), 164–169. https://doi.org/10.1111/j.1467-9280.1996.tb00350.x

Wynn, K. M. (1992). Addition and Subtraction in Human Infants. *Nature, 385*, 749–750.

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology, 25*(1), 103–108. https://doi.org/10.1348/026151005X90704

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*(1), 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x

Zuckermann, H., & Merton, R. K. (2010). Der Matthäus-Effekt in der Wissenschaft, II: Kumulativer Vorteil und der Symbolismus des intellektuellen Eigentums. *Berliner Journal für Soziologie, 20*(3), 285–308. https://doi.org/10.1007/s11609-010-0134-8

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement, 36*(1), 1–28. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x

Abbildungsverzeichnis

Tabellenverzeichnis